

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 02:00:14 ; Search time 82 Seconds
(without alignments)

12128.692 Million cell updates/sec

Title: US-09-830-762-1

Perfect score: 3243
Sequence: 1 ggagctcgtgtctcctggtg.....gtctgataaaaaaaaaa 3243

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCITUS.COMB.seq: *
6: /cgn2_6/ptodata/1/lna/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	202.8	6.3	214	4	US-09-222-575-24
C 2	65.4	2.0	7218	1	US-08-232-463-14
C 3	44	1.4	1529	3	US-09-189-760-5
C 4	44	1.4	1529	3	US-09-188-811-5
C 5	44	1.4	1529	4	US-09-514-422-5
C 6	43.2	1.3	2494	3	US-09-189-760-1
C 7	43.2	1.3	2494	4	US-09-514-422-1
C 8	41	1.3	17656	4	US-09-433-579-3
C 9	39	1.2	650	4	US-09-328-111-189
C 10	37	1.1	2615	1	US-08-072-281-1
C 11	37	1.1	2615	1	US-08-759-446-1
C 12	37	1.1	2615	1	US-09-027-998A-1
C 13	37	1.1	3050	2	US-09-031-442A-21
C 14	37	1.1	3050	4	US-09-258-377-21
C 15	37	1.1	6265	4	US-09-129-112-3
C 16	37	1.1	11485	4	US-09-410-464-9
C 17	36.6	1.1	16535	4	US-08-961-527-74
C 18	36.4	1.1	2951	1	US-08-413-118-104
C 19	36.4	1.1	2951	3	US-08-473-446-104
C 20	36.4	1.1	3000	1	US-08-220-151-1
C 21	36.4	1.1	3000	1	US-08-413-118-1
C 22	36.4	1.1	3000	3	US-08-473-446-1
C 23	36.2	1.1	936	3	US-08-910-501-3
C 24	36.2	1.1	936	3	US-09-398-550-3
C 25	36.2	1.1	939	4	US-08-910-501-1
C 26	36.2	1.1	939	4	US-09-398-550-1
C 27	35.8	1.1	414	1	US-08-377-687-48

28	35.8	1.1	414	1	US-08-777-192-48	Sequence 48, Appl
29	35.8	1.1	414	4	US-08-971-982-48	Sequence 48, Appl
30	35.8	1.1	414	4	US-09-077-951-19	Sequence 19, Appl
C 31	35.8	1.1	19124	2	US-08-487-826B-13	Sequence 13, Appl
C 32	35.6	1.1	18475	3	US-08-961-527-38	Sequence 38, Appl
C 33	35.6	1.1	51952	3	US-08-947-823-1	Sequence 1, Appl
C 34	35.4	1.1	3418	4	US-09-193-562D-29	Sequence 29, Appl
C 35	35.4	1.1	5181	4	US-08-257-573-10	Sequence 10, Appl
C 36	35.4	1.1	24183	4	US-08-943-731-3	Sequence 3, Appl
C 37	35.2	1.1	2960	3	US-08-913-842-3	Sequence 3, Appl
C 38	35.2	1.1	3231	3	US-08-632-806A-7	Sequence 7, Appl
C 39	35.2	1.1	5687	4	US-09-221-017B-368	Sequence 368, App
C 40	35.2	1.1	5852	1	US-07-867-106-2	Sequence 2, Appl
C 41	35.2	1.1	9115	4	US-07-753-520B-3	Sequence 3, Appl
C 42	35	1.1	812	4	US-09-227-357-115	Sequence 115, App
C 43	34.8	1.1	2570	2	US-09-056-075-2	Sequence 2, Appl
C 44	34.6	1.1	447	1	US-08-211-942-4	Sequence 4, Appl
C 45	34.6	1.1	732	1	US-08-211-942-6	Sequence 6, Appl

ALIGNMENTS

```

RESULT 1
US-09-222-575-24/c
; Sequence 24, Application US/09222575
; Patent No. 6387697
; GENERAL INFORMATION:
; APPLICANT: Yuguo, Jiang
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 210121.470
; CURRENT APPLICATION NUMBER: US/08/222,575
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 214
; TYPE: DNA
; ORGANISM: Human
US-09-222-575-24

Query Match
Best Local Similarity 96.7%; Score 202.8; DB 4; Length 214;
Matches 207; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2593 ATGATGTCATGACAGTGTGTTGACAAATGTCACAGCCCGTAACCTTGTATTTT 2652
DB 214 AAGATGTCATGACAGTGTGTTGACAAATGTCACAGCCCGTAACCTTGTATTTT 155
QY 2653 AGTGCATGACAGCTATACCTGATTTGTGACATGACACCGAGATATGACAGTCA 2712
DB 154 AGTGCATGACAGCTATACCTGATTTGTGACATGACACCGAGATATGACAGTCA 95
QY 2713 CAAGATTACTATGACAGCAATGACACCTTTGACAGCATTCCTTGAGAGGATTCGA 2772
DB 94 CAAGATTACTATGACAGCAATGACACCTTTGACAGCATTCCTTGAGAGGATTCGA 35
QY 2773 AGTGCAGCCCTGTTGATTAATCTACATCATTT 2806
DB 34 AGTGCAGCCCTGTTGATTAATTTTACATCATTT 1

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.

```

1 APPLICANT: FALKNER, F. G.
2 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
3 NUMBER OF SEQUENCES: 52
4 CORRESPONDENCE ADDRESSES:
5 ADDRESSEE: Foley & Lardner
6 STREET: 1800 Diagonal Road, Suite 500
7 CITY: Alexandria
8 STATE: VA
9 COUNTRY: USA
10 ZIP: 22313-0299
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC-DOS/MS-DOS
15 SOFTWARE: PatentIn Release #1.0, Version #1.25
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/08/232,463
18 FILING DATE:
19 CLASSIFICATION: 435
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: US/07/935,313
22 FILING DATE:
23 APPLICATION NUMBER: EP 91 114 300.6
24 FILING DATE: 26-AUG-1991
25 ATTORNEY/AGENT INFORMATION:
26 NAME: BENT, Stephen A.
27 REGISTRATION NUMBER: 29,768
28 REFERENCE/DOCKET NUMBER: 30472/114 IMKU
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (703)836-9300
31 TELEFAX: (703)683-4109
32 TELE: 899149
33 INFORMATION FOR SEQ ID NO: 14:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 7218 base pairs
36 TYPE: nucleic acid
37 STRANDEDNESS: single
38 TOPOLOGY: linear
39 IMMEDIATE SOURCE:
40 CLONE: PT9pt-F15
41 US-08-232-463-14

	Query Match	2.0%	Score 65.4	DB 1,	Length 7218:	
	Best Local Similarity	6.4%:	Pred. No. 5.2e-08:			
	Matches 27:	Conservative 230:	Mismatches 166:	Indels 0:	Gaps 0:	
QY	1945 ATCACAGAAATTATCAGCCGACGAAGAAGACAGATTAGAGATTGAAGAGACCCTGGCGTGCAC		:::: :::: :::: :::: :::: :::: ::::			2004
Db	1451 ATGACAAGATTGTGTACRRR					1392
OY	2005 CTCATTGAGCTGCTTATGAACATTGAAGTATGCTTTGTAAGTCACAGAGAAATGGCA					2064
Db	1391 RR					1332
OY	2065 AGGAAAAGAGACGAGCGCTTGCGGCTCGTAGACAGTGGACTATTCAGACAACAAAA	:	: : : : : : : :			2124
Db	1331 RR					1272
OY	2125 ACCCAGACGGAATCAAGAAATAATTCATATTTCGAATGTCTGTGCTCCAACCTCAA					2184
Db	1271 RR					1212
OY	2185 GAGCTTCAGATCAAAAGAACAGTTGAAAAAATGTGGTTTCAGTTCTAAAGGGGAG					2244
Db	1211 RR					1152
OY	2245 AAGATAGCAATGAGGCTCTTATGGCTCCCTTCAAGAAACAAGAAATAATGAGAGAA					2304
Db	1151 RR					1092
OY	2305 AAACGACAGGCAACCTGTGAGCATATGAGCGTTTCACAGAAGTCCCATACAGTTCTGC					2364
Db	1091 RRRRRRRRRRRRRRRRRRRRRRAATGCGAAGCTCCCTCGACCTGCGACCCAAAGTTCGA					1032

```

OY 2365 AAT 2367
Db 1031 ATT 1029

RESULT 3
US-09-189-760-5
; Sequence 5, Application US/09189760
; Patent No. 6031078
; GENERAL INFORMATION:
; APPLICANT: Rhodadoust, Mehrian
; TITLE OF INVENTION: NOVEL MTX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNT-046CP2
; CURRENT APPLICATION NUMBER: US/09/189,760
; CURRENT FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; EARLIER APPLICATION NUMBER: 60/089,467
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: (PENDING)
; EARLIER FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1529
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(749)
US-09-189-760-5

Query Match 1.4%; Score 44; DB 3; Length 1529;
Best Local Similarity 48.3%; Pred. No. 0.023;
Matches 113; Conservative 3; Mismatches 118; Indels 0; Gaps 0;

OY 1562 TTTTGTGGTGAACATTTTAAACATTTGCCAAGACCATGTAAACACCGTAACTGATA 1621
Db 873 TTTTGTGCATTTCTCTAAAGAAAGGTGCCAAAGCTTTTGATTTGCTGCAGGTACTGAAA 932
OY 1622 AATTTGTGATCTTTCCACACAGATTTGGAGATATATTAAGCTTTTCATCTTGAATGGCAA 1681
Db 993 CAACACTGACATTTTAAAGAAATATGATTAATGAGACCTTTAAGATTTTAAATTCGA 992
OY 1682 AGAGTCCACAGTCTGACTGTGAACAATTAACAGTGTCCCCAGTCACACAGAGGAGA 1741
Db 993 AGGGATCCAAAGGTTTGTATTTATCTTATTGGGGAGACACTAACMWTTCAAAGAAAGCAG 1052
OY 1742 AAAGAGAAATGGGCTTGAAGCTAGACTTCCTCCATCATCATGATGGAGATTCGA 1795
Db 1053 CTGTCAACATTTGGTGGCCAGTGTCTATCAGATGAGTTAAACCTTTGATTTCA 1106

RESULT 4
US-09-188-811-5
; Sequence 5, Application US/09188811
; Patent No. 6037148
; GENERAL INFORMATION:
; APPLICANT: Rhodadoust, Mehrian
; TITLE OF INVENTION: NOVEL MTX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNT-046CP
; CURRENT APPLICATION NUMBER: US/09/188,811
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1529
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
;
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: (3)..(749)
;
US-09-188-811-5

```

Query Match	1.4%;	Score 44;	DB 3;	Length 1529;
Best Local Similarity	48.3%;	Pred. NO. 0.023;		
Matches 113; Conservative	3;	Mismatches 118;	Indels 0;	Gaps 0;

QY	1562	TTTTGGTGGATGAAGTTTTTAACTTTGGCCAAAGACATGTAATAACACCAAGTTAACTGTA	1621
Db	873	TTTTGTGCATTTCTCTAAAGAAAGGTGCCAAACCTTTTGATTTGCTGCACGGTAACTGTA	932
QY	1622	AATTTTGATCTTTCCACAGATTTTGAGATATATTAAGGCTTTTCAGTTCTGTAATGGCA	1681
Db	933	CAAACTAGCATTTTAAAAATTAAGATTTAATGGAAGCCTTTAAGATTTTAATTCGA	992
QY	1682	AGAGGTCGAAGATGCTGAGTTTGAACAATTACAGTGTCCCGCAGTCAACACGAGAGGAGA	1741
Db	993	AGGATCCAAAGGTTCTGTATTATCTTATTGCGGAGACACTAACMMTTCAAGAAGCAGG	1052
QY	1742	AAAGAGAAATGGGGCTTGAAGCTAAGATTCGTCGCCATCAATGATGGGATTC	1795
Db	1053	CTGTGAACATTTGGGTGCCACAGTCTCAAGATGACTTAAACCTTTGATTCTCA	1106

RESULT 5
US-09-514-422-5

```

? Sequence 5 Application US/09514422
? Patent No. 6291193
? GENERAL INFORMATION:
? APPLICANT: Khodadoust, Mehrian
? TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
? TITLE OF INVENTION: THEREFOR
? FILE REFERENCE: MNI-046CP2
? CURRENT APPLICATION NUMBER: US/09/514,422
? CURRENT FILING DATE: 2000-02-28
? PRIOR APPLICATION NUMBER: US/09/189,760
? PRIOR FILING DATE: 1998-11-10
? PRIOR APPLICATION NUMBER: 09/163,116
? PRIOR FILING DATE: 1998-09-29
? PRIOR APPLICATION NUMBER: 60/089,467
? PRIOR FILING DATE: 1998-06-16
? PRIOR APPLICATION NUMBER: (PENDING)
? PRIOR FILING DATE: 1998-11-09
? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 5
? LENGTH: 1529
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (3)..(749)
? OS-09-514-422-5

```

Query Match	1.48;	Score 44;	DB 4;	Length 1529;
Best Local Similarity	48.38;	Pred. No. 0.023;		
Matches 113; Conservative	3;	Mismatches 118;	Indels 0;	Gaps 0

QY	1562	TTTGTTGGTGAAGTTTACATTTGGCAAGACATGTAAACACCAAGTTAACTGTAA	1621
Db	873	TTTTGTGCAATTCCTCAAAAGAAAGGCGCAAAACCTTTTGAATGGTGCAGGTAATGAA	932
QY	1622	AATTGTCATCTTCCACAGATTTGGAGATATATAAGCGTTTCAGTCTGAAATGGCAA	1681
Db	933	CAAACTGACATTTTAAAAAATTATAGCATTTAATGGAAGCGCTTAAAGATTTTAAATTCGA	992
QY	1682	AGAGGTCAAGATGCTAGCTGTTGAACAATTAACAGTGTCCCGCAGTCAACACAGACGAGAA	1741
Db	993	AGGATCCAAAGGTTCTGTATTTATCTTTATTTGGGACACACTAATCMTTTAAAGAAAGCAG	1051

Dy 1742 AAGAGAAATGSGCTTGAAGCTAGATCTCCGCCATCATCTGATGGATTCA 1795
| | |||| | | | | | |
Db 1053 CTGTACACATTGGGTGCCCAGTGCTTCAGATGAGTTAAAACCTTTGATTTCTCA 1106

RESULT 6
US-09-189-760-1
; Sequence 1, Application US/09189760

Query Match	1.3%	Score 43.2	DB 3	Length 2494
Best Local Similarity	47.4%	Pred. No. 0.05		
Matches 111	Conservative 5	Mismatches 118	Indels 0	Gaps 0

QY	1562	TTTGTGTGATGAAAGTTTAACTTTTGGCCAAAGACCATTAACACCGCATTAACGTAA	1622
		:	
Db	1838	KTTTGTGCATCTCTCTAAAGAAGGTGCCAAGCCTTTTGTGATGCTGCGAGTAACCTGAA	1897
QY	1622	AATTTTGATCTTTCCCAACAGATTTGTGAGATATATFAGGCTTTACGTCTCGAATGGCAA	1681
		: :	
Db	1898	CAAACTCAGACATTTTAAATAATTAAGATTAATGAGAGCTTTAAGATTTTAAATTTGCA	1955
QY	1682	AGAGGTCCAAGATCTGAGCTTTGAAACAATTACAGTGTCCGCCAGTCAACACAGAGAGAGA	1741
		: :	
Db	1958	AGGAGTCAAGGTTCTGTATTTATTTCTTATTGTGGGAGACACTAACMMTTCAAAGAGCAGG	2017
QY	1742	AAAGAGAAATGGGCTTGAAGCTAGATCTCTCGTCCATCAATCTAGTGGATTTCA	1795
		:	
Db	2018	CTGTGAACATTTGGGTGGCCACAGTCTCTGATGATGAATTAACACCTTTGATTTTCA	2071

RESULT 7
US-09-514-422-1
: Sequence 1, Application US/09514422

Patent No. 6291193
 GENERAL INFORMATION:
 APPLICANT: Rhododoust, Mehran
 TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 TITLE OF INVENTION: THEREFOR
 FILE REFERENCE: MT-046CP2
 CURRENT APPLICATION NUMBER: US/09/514,422
 CURRENT FILING DATE: 2000-02-28
 PRIOR APPLICATION NUMBER: US/09/189,760
 PRIOR FILING DATE: 1998-11-10
 PRIOR APPLICATION NUMBER: 09/163,116
 PRIOR FILING DATE: 1998-03-29
 PRIOR APPLICATION NUMBER: 60/089,467
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: (PENDING)
 PRIOR FILING DATE: 1998-11-09

NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 1
LENGTH: 2494
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(1714)
US-09-514-422-1

Query Match 1.3%; Score 43.2; DB 4; Length 2494;
Best Local Similarity 47.4%; Pred. No. 0.05;
Matches 111; Conservative 5; Mismatches 118; Indels 0; Gaps 0;

OY 1562 TTTTGTGATGATTTTAACTTTGCAAGACATGTAACACACCGTTAACTGTA 1621
DB 1838 KTKKGTGCAATTCCTAAAGAGAGTCCAAAGCTTTTGATGCTGCAAGTAACTGAAA 1897
OY 1622 AATTTGTGATCTTTCCACAGATTTGAGATATATAGAGCTTTCAGTCTGTAAGTGC 1681
DB 1898 CAACCTGACATTTTAAAGATTTATGATTTAGAGCTTTTAAAGATTTTAAATTCGA 1957
OY 1682 AGAGTCCAAAGTGTGATTTGAACATTTACAGTGTCCCGACATCAACACAGAGAGAGA 1741
DB 1958 AGGATCCAAAGTGTGATTTATCTTATTTGGGAGACACTAACMNTTCAAGAGAGCAG 2017
OY 1742 AAAGAGAAATAGGCTTGAAGCTAGATCTCCATCATCATGATGGATGCA 1795
DB 2018 CTGTGACATTTGGTGGCCAGTGTATCAGATGAGTTAAACCTTTGATTCICA 2071

RESULT 8
US-09-433-579-3
Sequence 3, Application US/09433579
Patent No. 644877

GENERAL INFORMATION:
APPLICANT: Rotmann, William H.
TITLE OF INVENTION: LSAG Gene
FILE REFERENCE: LSAG Gene
CURRENT APPLICATION NUMBER: US/09/433,579
CURRENT FILING DATE: 1999-11-04
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 3
LENGTH: 17656
TYPE: DNA
ORGANISM: Liquidambar styraciflua
US-09-433-579-3

Query Match 1.3%; Score 41; DB 4; Length 17656;
Best Local Similarity 49.8%; Pred. No. 0.6;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

OY 3035 TTGTTCTTAACTCTCCCTTACTAGAGAGATAATATTTTGCATATTAAGAATAATTT 3094
DB 9997 TGGTACTTATCTCTGATATGTAATAATATATATATATATATATATATATATA 10056
OY 3095 TCTAGTATATACGACGCTTTTATTTCTTAATGATGATATATAAATGTTAGA 3154
DB 10057 TACTGATATATGACACGCGCTTCACATGTAATTTTAAATATATATATAGGTAGA 10116
OY 3155 TAACAGATGATTTAGATTTTCCAGAGATATATTAAGAGCTTTAGGTATGAATAAT 3214
DB 10117 TAGATTAGTATTTTGTATATATATATAGTAGTATAGTATATACAGATATATAATAC 10176
OY 3215 ATCATCTTTGTCTGATTAATAAAAAA 3243
DB 10177 ATAAAGATTTATATTTTATATAAAGA 10205

RESULT 9
US-09-328-111-189/c

Sequence 189, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Calino, Theodore J.
APPLICANT: Dertl, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 189
LENGTH: 650
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(650)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-189

Query Match 1.2%; Score 39; DB 4; Length 650;
Best Local Similarity 49.2%; Pred. No. 0.37;
Matches 90; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

OY 2947 CTTAGTCCCTTATATGCTTTATGCTTGGGCTTGGAGTAGATACCAATGAAC 3006
DB 562 CCGTGTCTTTCTTCTCCCTTAATATTTGGCTTCAGATTTGCCGCCAAATCCAGA 503
OY 3007 ACTTCAGAGACCTTCCTCTCTTCAGTGTCTTAAATCTCTTACTAGAGAGATA 3066
DB 502 GTCTAATCTTCACCCCTTCCAGTGCAGTTAAATGACCCCTTCTAAAGTTAAG 443
OY 3067 AATATTTGCATATATGAAGAAATTTTCTAGTATATACGACGCTTTATTTCTA 3126
DB 442 CAAGGTTCACAGACAGAGAGATGATTCATTTGAAAAACCATAGATTTAGTTAATA 383
OY 3127 AAA 3129
DB 382 ATA 380

RESULT 10
US-08-072-281-1/c
Sequence 1, Application US/08072281
Patent No. 5495071
GENERAL INFORMATION:
APPLICANT: Fischhoff, David A.
APPLICANT: Fuchs, Roy L.
APPLICANT: Lavrik, Paul B.
APPLICANT: McPherson, Sylvia A.
APPLICANT: Perlak, Frederick J.
TITLE OF INVENTION: Insect Resistant Plants
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESS: Lawrence M. Lavin, Jr., Monsanto Co., BB4F
STREET: 700 Chesterfield Parkway No. 5495071th
CITY: St. Louis
STATE: Missouri
COUNTRY: United States of America


```
; LENGTH: 6265
; TYPE: DNA
; ORGANISM: Dolichos biflorus
; FEATURE:
; OTHER INFORMATION: genomic sequence of NBP46 (DB46)
; NAME/KEY: exon
; LOCATION: (633)..(944)
; NAME/KEY: intron
; LOCATION: (945)..(1022)
; NAME/KEY: exon
; LOCATION: (1023)..(1151)
; NAME/KEY: intron
; LOCATION: (1152)..(1559)
; NAME/KEY: exon
; LOCATION: (1560)..(1616)
; NAME/KEY: intron
; LOCATION: (1617)..(1697)
; NAME/KEY: exon
; LOCATION: (1698)..(1790)
US-09-129-112-3
```

Query Match 1.18; Score 37; DB 4; Length 6265;

Best Local Similarity 50.98; Pred. No. 4.6;

Matches 88; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```
OY 3048 TCCTTACTAGAGAGATAATATTTCATATATGAAGAATTTTCTAGTATATAAC 3107
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6145 TACTATACAAAAATAAAAAATAAAATTAATTAATCTTAATAATTTCTTAATTA 6086
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 3108 GCAGGCTTTTATTTCTAAATGATAGATATAAAATGTTAGATACAGATGATT 3167
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6085 TTAATATTAAATAATATAAATTATAAATAATATATCAAAATTAATATAAATAAT 6026
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 3168 TTAGATTTCAGAGATATATAAGTCTTAGTATGAATAATATCATC 3220
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6025 ATATATTGATTAATTAATTTAATAATTATTAATTCATTCATTAATTC 5973
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: February 7, 2003, 05:39:02

Job time : 204 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2003, 23:46:08 ; Search time 420 Seconds
(without alignments)
17388.644 Million cell updates/sec

Title: US-09-830-762-1
3243
Perfect score: 1 99gcctcgtctcctg99gtg.....gtctgattaaaaaaaaa 3243
Sequence:

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.GeneSeq_101002:*

- 1: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3243	100.0	3243	21	Human B-aggressive
2	2842.6	87.7	3016	24	Human Interferon-1
3	2543.8	78.4	2598	22	Human transcription
4	2533	78.1	2581	24	Human ovarian anti
5	1527.8	47.1	1695	22	Human polynucleoti
6	1526.2	47.1	1747	22	Human polynucleoti
7	1340	41.3	3024	21	Human B-aggressive
8	792	24.4	854	23	Human prostate exp
9	792	24.4	854	23	Human prostate exp

10	622.8	19.2	822	22	AAH72506	Human cervical can
11	609	18.8	820	23 <td>ABV30165</td> <td>Human prostate exp</td>	ABV30165	Human prostate exp
12	489.4	15.1	666	22 <td>AAH72659</td> <td>Human cervical can</td>	AAH72659	Human cervical can
13	463.2	14.3	516	22 <td>AAH70240</td> <td>Human cervical can</td>	AAH70240	Human cervical can
14	449.8	13.9	481	22 <td>AAH71670</td> <td>Human cervical can</td>	AAH71670	Human cervical can
15	424.2	13.1	668	24 <td>ABO56473</td> <td>Human colon cancer</td>	ABO56473	Human colon cancer
16	414.8	12.8	457	22 <td>AAH68812</td> <td>Human cervical can</td>	AAH68812	Human cervical can
17	379.4	11.7	389	22 <td>AAH20002</td> <td>Human breast cancer</td>	AAH20002	Human breast cancer
18	362.4	11.2	411	22 <td>AAH10422</td> <td>Human breast cancer</td>	AAH10422	Human breast cancer
19	329.4	10.2	643	24 <td>ABO56407</td> <td>Human colon cancer</td>	ABO56407	Human colon cancer
20	317	9.8	342	24 <td>ABV13942</td> <td>Human prostate exp</td>	ABV13942	Human prostate exp
21	317	9.8	386	23 <td>ABV35051</td> <td>Human prostate exp</td>	ABV35051	Human prostate exp
22	317	9.8	386	23 <td>ABV35051</td> <td>Human prostate exp</td>	ABV35051	Human prostate exp
23	316	9.7	356	23 <td>ABV4773</td> <td>Human prostate exp</td>	ABV4773	Human prostate exp
24	299	9.2	299	24 <td>ABK45395</td> <td>CDNA encoding colo</td>	ABK45395	CDNA encoding colo
25	284	8.8	284	24 <td>ABL64271</td> <td>Stomach cancer rel</td>	ABL64271	Stomach cancer rel
26	280.4	8.6	524	22 <td>AAH70685</td> <td>Human cervical can</td>	AAH70685	Human cervical can
27	273	8.4	378	23 <td>ABV15538</td> <td>Human prostate exp</td>	ABV15538	Human prostate exp
28	259	8.0	363	23 <td>ABV36329</td> <td>Human prostate exp</td>	ABV36329	Human prostate exp
29	257	7.9	363	23 <td>ABV24062</td> <td>Human prostate exp</td>	ABV24062	Human prostate exp
30	257	7.9	363	23 <td>ABV29943</td> <td>Human prostate exp</td>	ABV29943	Human prostate exp
31	248.4	7.7	463	23 <td>ABV45661</td> <td>Human prostate exp</td>	ABV45661	Human prostate exp
32	246.8	7.6	260	24 <td>ABL62550</td> <td>Colon adenocarcino</td>	ABL62550	Colon adenocarcino
33	246.8	7.6	260	24 <td>ABL67274</td> <td>Thyroid cancer rel</td>	ABL67274	Thyroid cancer rel
34	243	7.5	360	23 <td>ABV45360</td> <td>Human prostate can</td>	ABV45360	Human prostate can
35	209.8	6.5	387	22 <td>AAH71892</td> <td>Human breast can</td>	AAH71892	Human breast can
36	202.8	6.3	214	22 <td>AAH46970</td> <td>Human breast can</td>	AAH46970	Human breast can
37	202.8	6.3	214	22 <td>AAH71540</td> <td>Human breast can</td>	AAH71540	Human breast can
38	202.8	6.3	214	22 <td>ABK95006</td> <td>Human breast tumou</td>	ABK95006	Human breast tumou
39	183.4	5.7	343	24 <td>ABN16310</td> <td>Human OPRX polynuc</td>	ABN16310	Human OPRX polynuc
40	181.2	5.6	183	24 <td>ABL68001</td> <td>Human colon tumour</td>	ABL68001	Human colon tumour
41	168.4	5.2	172	20 <td>AAH40690</td> <td>Human secreted pro</td>	AAH40690	Human secreted pro
42	167.4	5.2	170	21 <td>AAC29138</td> <td>Human secreted pro</td>	AAC29138	Human secreted pro
43	162	5.0	317	23 <td>ABV15861</td> <td>Human prostate exp</td>	ABV15861	Human prostate exp
44	158.4	4.9	185	23 <td>ABV06369</td> <td>Human prostate exp</td>	ABV06369	Human prostate exp
45	140.4	4.3	2269	22 <td>AAH89896</td> <td>Human bone marrow</td>	AAH89896	Human bone marrow

ALIGNMENTS

RESULT 1
ID AAD00629 standard: CDNA: 3243 BP.
AAD00629

29-AUG-2000 (first entry)

Human B-aggressive lymphoma (BAL) CDNA.

Human: B-aggressive lymphoma; BAL; chromosome 3q21; tumour; malignancy;
differential expression; DB-CL; diffuse large B-cell lymphoma;
cytostatic; vaccine; treatment; diagnosis; non-Hodgkin's lymphoma;
cellular adhesion; sarcoma; carcinoma; myeloma; ss.

Homo sapiens.

Key	Location/Qualifiers
5'UTR	1..228
FT	/*tag= a
FT	229..2793
FT	/*tag= b
FT	/*product= "BAL protein"
FT	229..2790
FT	/*tag= c
FT	/*note= "This region is specifically claimed in Claim 1 as SEQ ID NO.3"
FT	2794..3243
FT	/*tag= c
FT	278..382
FT	/*tag= d
FT	/*note= "Alternatively spliced sequence"

FT misc-feature 2528..3132
FT /*tag= e
FT /note= "This region shows 98% identity to Soares
FT pregnant uterus NBHPU human cDNA clone 502921."
XX
XX MO200026231-A1.
XX
XX 11-MAY-2000.
XX
XX 29-OCT-1999; 99MO-US25439.
XX
XX 29-OCT-1998; 98US-0106383.
XX 30-OCT-1998; 98US-0106448.
XX
XX (DAND) DANA FARBER CANCER INST INC.
XX
XX Shipp M, Aguilar R, Yakushiji J Y;
XX WPI; 2000-365563/31.
XX P-PSDB; AAY71082.
XX
XX New B-aggressive lymphoma nucleic acid for identifying cells exhibiting
PT or predisposed to malignancies such as lymphoma, sarcoma, carcinoma and
PT myeloma
XX
XX Claim 1; Fig 1; 151pp; English.
XX
XX The present sequence is a cDNA encoding B-aggressive lymphoma (BAL)
CC protein obtained by screening two human cDNA libraries derived from
CC anti-immunoglobulin activated splenocytes and the Raji Burkitts lymphoma
CC cell line cloned into pCDM8. The BAL locus is mapped to chromosome 3q21.
CC BAL was identified to be differentially expressed in DIB-CL (Diffuse
CC large B-cell lymphoma), the most common non-Hodgkin's lymphoma, by
CC differential display technique. It was found to be more abundant in
CC tumours from patients with high risk fatal DIB-CL disease than low risk
CC cured disease. BAL transcripts are most abundant in lymphoid organs e.g.
CC spleen, foetal liver and peripheral blood, and several non-haematopoietic
CC organs e.g. heart and skeletal muscle. BAL is involved in modulation of
CC cellular adhesion and aggressiveness/severity of malignancy such as
CC DIB-CL. BAL nucleic acid and protein are useful for identifying cells
CC exhibiting or predisposed to malignancies such as lymphoma, sarcoma,
CC carcinoma and myeloma. BAL proteins are useful as immunogens to raise
CC anti-BAL antibodies. BAL modulators or BAL molecules can be used for
CC prophylactic and therapeutic treatment of a subject susceptible to or
CC having a disorder associated with aberrant BAL expression or activity,
CC such as non-Hodgkin's lymphoma. It is also useful in vaccine preparation.
XX
XX Sequence 3243 BP; 990 A; 647 C; 755 G; 851 T; 0 other;
SQ
Query Match 100.0%; Score 3243; DB 21; Length 3243;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 TTGTTTCAGAAAGTCTTTGCTCAGATCTTCTCAGTGGAGAAAGGGGAATACAGAGAA 350
DB 301 TTGTTTCAGAAAGTCTTTGCTCAGATCTTCTCAGTGGAGAAAGGGGAATACAGAGAA 350
QY 361 TGTCTCCCTPACAAAGTCTCAGAGACTGTGCTCTTGGAGAAACATTAATGTGGCAATT 420
DB 361 TGTCTCCCTPACAAAGTCTCAGAGACTGTGCTCTTGGAGAAACATTAATGTGGCAATT 420
QY 421 CCCATTAAACCAATGACTCAAAATTTTAAAAAATTAATAGCGTCAAGCTGTGAAGTC 480
DB 421 CCCATTAAACCAATGACTCAAAATTTTAAAAAATTAATAGCGTCAAGCTGTGAAGTC 480
QY 481 CTCACAGATAAGTTGGCTATCTGTACCCCTGGTCTCTCCAGTTCAGAGAACACAGC 540
DB 481 CTCACAGATAAGTTGGCTATCTGTACCCCTGGTCTCTCCAGTTCAGAGAACACAGC 540
QY 541 AAATCTGTGAAGTGTGCAAAAAATGCTGACCTGCTAGTGTAGAGTATAGCTGTGAAA 600
DB 541 AAATCTGTGAAGTGTGCAAAAAATGCTGACCTGCTAGTGTAGAGTATAGCTGTGAAA 600
QY 601 GATGACCTCACACACATGCTGTTGATGCTGTGGTAATGACCAATGAAGATCTTCTG 650
DB 601 GATGACCTCACACACATGCTGTTGATGCTGTGGTAATGACCAATGAAGATCTTCTG 650
QY 661 CATGGGGAGGCGCTGGCCCTGGCTGTAAAGCTGGTGATTTGAATCAAGAAAG 720
DB 661 CATGGGGAGGCGCTGGCCCTGGCTGTAAAGCTGGTGATTTGAATCAAGAAAG 720
QY 721 AGCAAAAGTTTGTGGCAATATGTAAGTGTGACGCTGTAGATAGCTGTACAGGGA 780
DB 721 AGCAAAAGTTTGTGGCAATATGTAAGTGTGACGCTGTAGATAGCTGTACAGGGA 780
QY 781 GCAGGAGGCTTCCTGCAAAACAGATCATGCTGTTGGGCTCGGTGATGATGATG 840
DB 781 GCAGGAGGCTTCCTGCAAAACAGATCATGCTGTTGGGCTCGGTGATGATGATG 840
QY 841 GATTAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 841 GATTAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 901 ATGATTAATAATATCTGATTAAGACATGATGATGATGATGATGATGATGATGATG 960
DB 901 ATGATTAATAATATCTGATTAAGACATGATGATGATGATGATGATGATGATGATG 960
QY 961 CAGTTCCTCTGATTTGTGTACAAAGACTATTTGTAGAGCTATCCGGGTTAGTTTGCA 1020
DB 961 CAGTTCCTCTGATTTGTGTACAAAGACTATTTGTAGAGCTATCCGGGTTAGTTTGCA 1020
QY 1021 GGGAGGCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1021 GGGAGGCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 GTTGTGCTCTTAAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 1140
DB 1081 GTTGTGCTCTTAAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 1140
QY 1141 ACCACCCCTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
DB 1141 ACCACCCCTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 CACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
DB 1201 CACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 GTTGGACCTGTGCAAAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1320
DB 1261 GTTGGACCTGTGCAAAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1320
QY 1321 CTTGACCAAAAGGCTTAAAGCTTCAAGGCTCCAGTGTGATGATGATGATGATGATG 1380
DB 1321 CTTGACCAAAAGGCTTAAAGCTTCAAGGCTCCAGTGTGATGATGATGATGATGATG 1380
QY 1381 AACTTGTCTGTAAATATATATACATGATGATGATGATGATGATGATGATGATGATG 1440

```

|||||
Db 1381 AACCTGTTCTGTAATAATATATACATGCTACTGTGTCATTCAGAAATTTCTTAAACCTCAG 1440
OY 1441 AATATTAACATGCATGATGAGAGTGTGTTGAAAAATGCATGTAGCAAAAATATACCTCC 1500
Db 1441 AATATTAACATGCATGATGAGAGTGTGTTGAAAAATGCATGTAGCAAAAATATACCTCC 1500
OY 1501 AATTCCTTTCCTGCTGCTGGAGCTGGAAACATGAAATATAGAGAAACACACAGCAGAG 1560
Db 1501 AATTCCTTTCCTGCTGCTGGAGCTGGAAACATGAAATATAGAGAAACACACAGCAGAG 1560
OY 1561 AATTTGTTTGTATGAAGTTTAAACATTTGGCAAAGACCATGTAAACACAGTTAACTGTA 1620
Db 1561 AATTTGTTTGTATGAAGTTTAAACATTTGGCAAAGACCATGTAAACACAGTTAACTGTA 1620
OY 1621 AATTTGTTTGTATGCTTCCACACAGATTTGGAGATATATAGGCTTTTCAGTTCTGAAATGCA 1680
Db 1621 AATTTGTTTGTATGCTTCCACACAGATTTGGAGATATATAGGCTTTTCAGTTCTGAAATGCA 1680
OY 1681 AAGAGGTCACAGATGCTGAGTTTGAACATTTACAGTGTGCTCCAGTCACACAGAGAGAG 1740
Db 1681 AAGAGGTCACAGATGCTGAGTTTGAACATTTACAGTGTGCTCCAGTCACACAGAGAGAG 1740
OY 1741 AAAAGAGAAATGGGCTTGAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db 1741 AAAAGAGAAATGGGCTTGAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
OY 1801 GAGAGATATATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1860
Db 1801 GAGAGATATATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1860
OY 1861 ATCATTTGACATATATCATATTTCTGCTTGGAGAGAAAGAACATGACATTTTGTCTGAG 1920
Db 1861 ATCATTTGACATATATCATATTTCTGCTTGGAGAGAAAGAACATGACATTTTGTCTGAG 1920
OY 1921 CTTGAGAAAACTTCAAGTGTCTCATCACAGAAATTTATCAGCCACAGAGAGAGAGTTA 1980
Db 1921 CTTGAGAAAACTTCAAGTGTCTCATCACAGAAATTTATCAGCCACAGAGAGAGAGTTA 1980
OY 1981 GAGATTGAAGGAGCCCGGGCTGACCTCATTTGAGGCTGTTATGAACTTGAAGATATGCTT 2040
Db 1981 GAGATTGAAGGAGCCCGGGCTGACCTCATTTGAGGCTGTTATGAACTTGAAGATATGCTT 2040
OY 2041 TGTAAAGTACAGAGAGAAATGGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Db 2041 TGTAAAGTACAGAGAGAAATGGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
OY 2101 CAGTGTACTATTTCAGACACAAAAAACCAAGAGAGAAATGAAGAAATATCATATTTCTG 2160
Db 2101 CAGTGTACTATTTCAGACACAAAAAACCAAGAGAGAAATGAAGAAATATCATATTTCTG 2160
OY 2161 AATATGCTGTGCTGCTCCAACTCAAGAGCTTCTAGATCAAAAGAAACAGTTTGAAGAAATGT 2220
Db 2161 AATATGCTGTGCTGCTCCAACTCAAGAGCTTCTAGATCAAAAGAAACAGTTTGAAGAAATGT 2220
OY 2221 GGTTCGAGGTTCTAAAGTGTGAGAGATAGACAAATGAGTCTCTTATGGCTGCTTCAA 2280
Db 2221 GGTTCGAGGTTCTAAAGTGTGAGAGATAGACAAATGAGTCTCTTATGGCTGCTTCAA 2280
OY 2281 AGAAAGAGAGAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
Db 2281 AGAAAGAGAGAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
OY 2341 CAGCAAGTCCCATACAGTTCATGCTGATGTGATGAGAGAGTGTGCTTCAAGAAATGTAC 2400
Db 2341 CAGCAAGTCCCATACAGTTCATGCTGATGTGATGAGAGAGTGTGCTTCAAGAAATGTAC 2400
OY 2401 TCGACACCTTGGGATCCAAATACGAGGCTGGCATATATCTTCCAAAGAACCTCAAAAC 2460
Db 2401 TCGACACCTTGGGATCCAAATACGAGGCTGGCATATATCTTCCAAAGAACCTCAAAAC 2460
OY 2461 CTGGCAGAGAGAGCCAAAGAAATCTCTGCTGACATAGCTGATGTGTTGAGGCT 2520
|||||

```

```

Db 2461 CTGGCAGAGAGAGCCAAAGAAATCTCTGCTGACATATAGCTATGTGTTGAGGCT 2520
OY 2521 GAAGTACACAGAGGCTTCTTCCAGGAGACATCCGTTAAATATGTTCGCCACACAG 2580
Db 2521 GAAGTACACAGAGGCTTCTTCCAGGAGACATCCGTTAAATATGTTCGCCACACAG 2580
OY 2581 AGTCCGTGGAGATATAGATGCTATGACAGTGTGTTGACAAATGCTCCAGCCCTGAAC 2640
Db 2581 AGTCCGTGGAGATATAGATGCTATGACAGTGTGTTGACAAATGCTCCAGCCCTGAAC 2640
OY 2641 TTTGTTATTTTATAGTGACATGACAGGCTATACCTCAGTATTTGTGACATGACACAG 2700
Db 2641 TTTGTTATTTTATAGTGACATGACAGGCTATACCTCAGTATTTGTGACATGACACAG 2700
OY 2701 TATGTACATGACAAATATATCTATGACAGAGCAATGAGACCCCTTGGACAGATCCT 2760
Db 2701 TATGTACATGACAAATATATCTATGACAGAGCAATGAGACCCCTTGGACAGATCCT 2760
OY 2761 AGGGGATTTGCAAGTGGCAGCCCTGTTGATTTATCTCTACATCATTTTAAACGCTGTAT 2820
Db 2761 AGGGGATTTGCAAGTGGCAGCCCTGTTGATTTATCTCTACATCATTTTAAACGCTGTAT 2820
OY 2821 GGCTTTACCTTGGGTGAACTAACCAAAATATGACATGATGCTCAAGAGTGGCTTGA 2880
Db 2821 GGCTTTACCTTGGGTGAACTAACCAAAATATGACATGATGCTCAAGAGTGGCTTGA 2880
OY 2881 ATATATCCCATGGGTTATCTGATATGACATGAGGTTATTTGAAAGACTATGCCATATC 2940
Db 2881 ATATATCCCATGGGTTATCTGATATGACATGAGGTTATTTGAAAGACTATGCCATATC 2940
OY 2941 TAGCATCTTATGACCTTTTATCTGCTTTTATGCTTGGGTTGGGATAGATATACAA 3000
Db 2941 TAGCATCTTATGACCTTTTATCTGCTTTTATGCTTGGGTTGGGATAGATATACAA 3000
OY 3001 TGAACACCTTTCAGACCTTCTCTCTGAGAGTGTCTTAAATCTCTTACTAGAG 3060
Db 3001 TGAACACCTTTCAGACCTTCTCTCTGAGAGTGTCTTAAATCTCTTACTAGAG 3060
OY 3061 GAGATTAATTTTGTGATTAATGAAGAAATTTTCTACTATATTAACGAGGCTTTAT 3120
Db 3061 GAGATTAATTTTGTGATTAATGAAGAAATTTTCTACTATATTAACGAGGCTTTAT 3120
OY 3121 TTTCTAAATGATAGATATTAATTAATGTTAGATTAACGAATGATTTTATCCAG 3180
Db 3121 TTTCTAAATGATAGATATTAATTAATGTTAGATTAACGAATGATTTTATCCAG 3180
OY 3181 AGAATATTTAAAGTCTTTAGGATGAAATTAATCATCTTTGTCTGATTTAAAAAAA 3240
Db 3181 AGAATATTTAAAGTCTTTAGGATGAAATTAATCATCTTTGTCTGATTTAAAAAAA 3240
OY 3241 AAA 3243
Db 3241 AAA 3243

RESULT 2
ABA93022
ID ABA93022 standard. cDNA; 3016 BP.
XX
XX ABA93022;
XX
XX 10-APR-2002 (first entry)
XX
XX Human interferon-induced tetraspan (IT) protein encoding cDNA.
XX
XX Human: interferon-induced tetraspan; IT; cytosolic; anti-HIV;
XX immunosuppressive; antiinflammatory gene therapy; autoimmune disease;
XX inflammatory disease; cardiovascular disease; metabolic disease;
XX cancer growth inhibition; metastasis; gene; ss.
OS Homo sapiens.
XX
XX key Location/Qualifiers
FH

```

FT	CDS	123..2582	
FT		/tag=a	
FT		/product="interferon-induced tetraspan protein"	
FT		/note="this CDS encodes the protein given in ABB05392"	
FT	CDS	123..2456	
FT		/tag=b	
FT		/partial	
FT		/product="interferon-induced tetraspan protein"	
FT		/transl_except= (pos:1371..1373,aa:117)	
FT		/transl_except= (pos:2427..2432,aa:117)	
FT		/note="this CDS encodes the protein given in ABB05391; no stop codon given"	
XX			
PN		WO200196399-A2.	
XX			
PD		20-DEC-2001.	
XX			
XX			
PF		15-JUN-2001; 2001WO-US19305.	
XX			
PR		15-JUN-2000; 2000US-211565P.	
XX			
PA		(CURA-) CURAGEN CORP.	
PA		(BIOJ) BIOGEN INC.	
XX			
PI		Peyman JA, Da Silva A, Hockman P;	
XX			
DR		WPI; 2002-098056/13.	
DR		P-PSDB; ABB05391, ABB05392.	
XX			
PT		Novel isolated interferon-induced tetraspan polypeptide, useful for	
PT		treating autoimmune, inflammatory, cardiovascular and metabolic	
PT		diseases, and for inhibiting cancer growth and metastasis -	
XX			
PS		Claim 9; Page 6-7; 105pp; English.	
-XX			
CC		The present sequence encodes human interferon-induced tetraspan (IIT)	
CC		protein (I). (I) has cytosolic, anti-HIV, immunosuppressive and	
CC		antiinflammatory activities. The polynucleotide (II) encoding (I) can be	
CC		used in gene therapy. (I) and (II) can be used for treating or preventing	
CC		a pathology associated with (I) in a subject, preferably human, or for	
CC		the manufacture of a medicament for treating a syndrome associated with a	
CC		human disease selected from a pathology associated with (I). (I) and (II)	
CC		can be used for treating autoimmune diseases, inflammatory diseases,	
CC		cardiovascular diseases, metabolic diseases, and for inhibiting cancer	
CC		growth and metastasis. (I) or (II) can be used as diagnostic and/or	
CC		prognostic markers, in gene therapy, as research tools, and for tissue	
CC		regeneration in vitro and in vivo.	
XX			
SQ		Sequence 3016 BP; 932 A; 589 C; 696 G; 799 T; 0 other;	
	Query Match	87.7%; Score 2842.6; DB 24; Length 3016;	
	Best Local Similarity	96.3%; Pred. No. 0;	
	Matches 2636; Conservative	0; Mismatches 9; Indels 105; Gaps	14
QY	151 CATATCTGGAACACTACATGCTATGCTTTGGAGCGCAAAAGGAATTAACATTTAAAGACT	210	
DB	45 CAGGCTGTGAAACTACAGCTATGCTTTGAACGCGCAAAAGGAATTAACATTTAAAGACT	104	
QY	211 CCCCCGGGACCTGAGAGATGACATTTTCCATGAGTGCGCGGAGACGACACTTACATGAA	270	
DB	105 CCCCCGGGACCTGAGAGATGACATTTTCCATGAGTGCGCGGAGACGACACTTACATGAA	164	
QY	271 AATATCAGTAGATTAACCTCGCTCACCTCTGTTTCACAAAGTCTTTCTCAGATCTTT	330	
DB	165 AAA-----	167	
QY	331 CCTCAGTGAGAAAGGGAATACAGAGAATGTCTCCCTACAGATGCTCAGAGACTGCT	390	
DB	168 -----TCAGAGACTGCT	179	
QY	391 GCTCTTGGAGAAACTATATAGTTGGCAAAATCCCATTAACCAAAATGACTTCAAAATTTTA	450	
DB	180 GCTCTTGGAGAAACTATATAGTTGGCAAAATCCCATTAACCAAAATGACTTCAAAATTTTA	239	

OY	451	AAAAATAATGAGCTCAGCTGTGTGAAGTCTCTCCACAATAATGATTGGCTGTATCTCTAC	510
Db	240	AAAAATATGAGCTCAGCTGTGTGAAGTCTCTCCACAATAATGATTGGCTGTATCTCTAC	299
OY	511	CTGGCTCTCCAGTTCCAGGAAAGCAACAGCAATCTCTGAAAGTGTGCAAAATGCG	570
Db	300	CTGGCTCTCCAGTTCCAGGAAAGCAACAAATAATCTCTGAAAGTGTGCAAAATGCG	359
OY	571	ACTCCTAGATAGAGTTATATGATCTGTGAAAGATGAGACCTCACACATCTCTGATGCT	630
Db	360	ACTCCTAGATAGAGTTATATGATCTGTGAAAGATGAGACCTCACACATCTCTGATGCT	419
OY	631	GTGTGAATGACCAATGAAGATCTTCTGCAATGGGGAGGCTTGCCCTGGCCCTGGTA	690
Db	420	GTGTGAATGACCAATGAAGATCTTCTGCAATGGGGAGGCTTGCCCTGGCCCTGGTA	479
OY	691	AAAGCTGATGATTTGAAATCCAAAGAAAGAGCAACAGATTGTCGCCGATATGTGTA	750
Db	480	AAAGCTGATGATTTGAAATCCAAAGAAAGAGCAACAGATTGTCGCCGATATGTGTA	539
OY	751	GTGTGACGTGTGATGATGCTGTACAGGGAGGAGGAGGCTTCCCTGCAACAGATATC	810
Db	540	GTGTGACGTGTGATGATGCTGTACAGGGAGGAGGAGGCTTCCCTGCAACAGATATC	599
OY	811	CATGCTGTTGGCCCTCGGTGATGGAATGGAATGAACAGGAGATGATCTGGAAGCTGCAG	870
Db	600	CATGCTGTTGGCCCTCGGTGATGGAATGGAATGAACAGGAGATGATCTGGAAGCTGCAG	659
OY	871	AGGGCCATTGTAAGTATTTCTGATATGTATCTATATAAAATATCTACATTTAAGACATG	930
Db	660	AGGGCCATTGTAAGTATTTCTGATATGTATCTATATAAAATATCTACATTTAAGACATG	719
OY	931	GCAATTCACGCTTGAGACTCTGGGATTTTCACTTCCCTGATTTTGTATCAAAAGACT	990
Db	720	GCAATTCACGCTTGAGACTCTGGGATTTTCACTTCCCTGATTTTGTATCAAAAGACT	779
OY	991	ATTGTAGAGACTATCCGGGTTAGTTTGTGCAAGGAGGCCAATGATGATTTGGAAGA	1050
Db	780	ATTGTAGAGACTATCCGGGTTAGTTTGTGCAAGGAGGCCAATGATGATTTGGAAGA	839
OY	1051	ATTCACTGGTGAGCAATGAGAGCCTACTGTGTGCTTTAAGCTCTTCAACAATTC	1110
Db	840	ATTCACTGGTGAGCAATGAGAGCCTACTGTGTGCTTTAAGCTCTTCAACAATTC	899
OY	1111	ATTCGAGGGAAGTGTGAGCGGCAAGAAACCAACCCTCTTTCATGACATGGTCTG	1170
Db	900	ATTCGAGGGAAGTGTGAGCGGCAAGAAACCAACCCTCTTTCATGACATGGTCTG	959
OY	1171	AACAACCGAGCCCTCCAGATTGTTCACAGGGCACATTTGAATGGCAGACGAGATGTAAT	1230
Db	960	AACAACCGAGCCCTCCAGATTGTTCACAGGGCACATTTGAATGGCAGACGAGATGTAAT	1019
OY	1231	GTTAATTCGTAANCCACATGATATTACAGTTGACCTGTGGCAAGTCAATTCACAA	1290
Db	1020	GTTAATTCGTAANCCACATGATATTACAGTTGACCTGTGGCAAGTCAATTCACAA	1079
OY	1291	CAACGAGAGTGTGAATGAANTGGAATTCCTTGCACAAAGGCTTAAACAGTTTCAAGG	1350
Db	1080	CAACGAGAGTGTGAATGAANTGGAATTCCTTGCACAAAGGCTTAAACAGTTTCAAGG	1139
OY	1351	TCCGAGTTGTGACGTGTACAAAAAGATTAACTGTTCTGTAAATATATATACCATGTA	1410
Db	1140	TCCGAGTTGTGACGTGTACAAAAAGATTAACTGTTCTGTAAATATATATACCATGTA	1199
OY	1411	CTGTGGCATTTGAGATTTCTTAAACCTCAGATATTAACATGCAATGAAGAGTGTG	1470
Db	1200	CTGTGGCATTTGAGATTTCTTAAACCTCAGATATTAACATGCAATGAAGAGTGTG	1259
OY	1471	GAATAATGATGTGAGCAAAATATATACATTTCTTCTCTGACCCTGTGGAGCTGGAAC	1530
Db	1260	GAATAATGATGTGAGCAAAATATATACATTTCTTCTCTGACCCTGTGGAGCTGGAAC	1319

QY 1531 ATGGAATTAAGAGAGAAACAGCAGAGATTGTTGATGAGTTTAACTTTGGC 1590
 DB 1320 ATGGAATTAAGAGAGAAACAGCAGAGATTGTTGATGAGTTTAACTTTGGC 1379
 QY 1591 AAAGACCATGTAACACCACTGTAACCTGTAACCTGTAACCTGTAACCTG 1650
 DB 1380 AAAGACCATGTAACACCACTGTAACCTGTAACCTGTAACCTGTAACCTG 1439
 QY 1651 AATATTAAGCTTTCAGTCTGTAATGGAAGAGTCCAAATCTGATGTAACAT 1710
 DB 1440 AATATTAAGCTTTCAGTCTGTAATGGAAGAGTCCAAATCTGATGTAACAT 1499
 QY 1711 TACAGTGTCCCTCAGTCAACAGAGAGAAAGAAATGAGCTTAACTGATCT 1770
 DB 1500 TACAGTGTCCCTCAGTCAACAGAGAGAAAGAAATGAGCTTAACTGATCT 1559
 QY 1771 CCTGCATCAATCTGATGAGATTCACGTAAGAGATGATAGGCCCACTGATC 1830
 DB 1560 CCTGCATCAATCTGATGAGATTCACGTAAGAGATGATAGGCCCACTGATC 1619
 QY 1831 CAAGATCTCAGTCTCAGAACCACTCATTTGAAATATCATATCTGTAACCT 1890
 DB 1620 CAAGATCTCAGTCTCAGAACCACTCATTTGAAATATCATATCTGTAACCT 1679
 QY 1891 GGGAGAAAGAGACATGACATTTTGTCTCAGTTCAGAAACTTCAAGTGTCTCACA 1950
 DB 1680 GGGAGAAAGAGACATGACATTTTGTCTCAGTTCAGAAACTTCAAGTGTCTCACA 1739
 QY 1951 GAAATATCAGCCAGAGAGAGAGAGATTAAGAGAGAGAGAGAGAGAGAGAG 2010
 DB 1740 GAAATATCAGCCAGAGAGAGAGAGATTAAGAGAGAGAGAGAGAGAGAGAG 1799
 QY 2011 GAGGTGTTATGACATTAAGATTAAGTCTTAAAGTCAAGAGAGAAATGCAAGAA 2070
 DB 1800 GAGGTGTTATGACATTAAGATTAAGTCTTAAAGTCAAGAGAGAAATGCAAGAA 1859
 QY 2071 AAGGAGCGAGGCTTGTGCGCTGTAGAGAGAGTACTTTCAGCAACAAACCCAA 2130
 DB 1860 AAGGAGCGAGGCTTGTGCGCTGTAGAGAGAGTACTTTCAGCAACAAACCCAA 1919
 QY 2131 GACGAATTAAGAAATATCATATTTCTGAATGCTCTGCTCCTCAACCTCAAGACT 2190
 DB 1920 GACGAATTAAGAAATATCATATTTCTGAATGCTCTGCTCCTCAACCTCAAGACT 1979
 QY 2191 CTAGATCAAAAGAAACAGTTGAAAAATGCTTGCAGATCTAAAGGTGGAAGATA 2250
 DB 1980 CTAGATCAAAAGAAACAGTTGAAAAATGCTTGCAGATCTAAAGGTGGAAGATA 2039
 QY 2251 GACATGAGTCTTATGCTCTTTCAAAGAGAAAGAAATGATGGAAGAAACCTG 2310
 DB 2040 GACATGAGTCTTATGCTCTTTCAAAGAGAAAGAAATGATGGAAGAAACCTG 2099
 QY 2311 CACAGCAACCTGTGAGCCTATGCTTTCAGCAAGTCCCATACAGTTCGCAATGTG 2370
 DB 2100 CACAGCAACCTGTGAGCCTATGCTTTCAGCAAGTCCCATACAGTTCGCAATGTG 2159
 QY 2371 GATGAGAGTGTGCTTCAAGAAATGATCTGACACCTTGGAGTCAAAATGAGAGGT 2430
 DB 2160 GATGAGAGTGTGCTTCAAGAAATGATCTGACACCTTGGAGTCAAAATGAGAGGT 2219
 QY 2431 GGCATATCTTACCAAGAACTCAAAACCTGCGAGAGAGAGAGAGAGAGAGAGAG 2490
 DB 2220 GGCATATCTTACCAAGAACTCAAAACCTGCGAGAGAGAGAGAGAGAGAGAGAG 2279
 QY 2491 GCAGATAGCTGATGATGTTGTTGAGGCTGAAGTACTACAGGCTTCTTTCAGGGA 2550
 DB 2280 GCAGATAGCTGATGATGTTGTTGAGGCTGAAGTACTACAGGCTTCTTTCAGGGA 2339
 QY 2551 CATCGTTAAATATGTTCCCAACACAGTCTGAGAGATATAGATGATGATGACAT 2610
 DB 2340 CATCGTTAAATATGTTCCCAACACAGTCTGAGAGATATAGATGATGATGACAT 2399
 QY 2611 GTGTTGACAAATGTCTCAGCCCTGAACCTTGTATTATTTAGTGCATGAGGCTATA 2670

DB 2400 GTGTTGACAAATGTCTCCAGCCCTGAACCTTGTATTATTTAGGCGATGAGCTATA 2459
 QY 2671 CCTCAGTATTTGTGACATGACCCAGAGATATGACAGTCAACAGATTACTCATCAGA 2730
 DB 2460 CCTCAGTATTTGTGACATGACCCAGAGATATGACAGTCAACAGATTACTCATCAGA 2519
 QY 2731 CCAATGAGCCCTTGCACAGATCTTGGAGGGATTCGCAAGTGGAGCCCTGTAT 2790
 DB 2520 CCAATGAGCCCTTGCACAGATCTTGGAGGGATTCGCAAGTGGAGCCCTGTAT 2579
 QY 2791 TATCTACATCATTTTAAACAGCTGTATGCTTACCTTGGTGTGAATCAACAAATAA 2850
 DB 2580 TATCTACATCATTTTAAACAGCTGTATGCTTACCTTGGTGTGAATCAACAAATAA 2639
 QY 2851 TGACCATGATGCTCAAGAGAGTGTGATATATCCATGAGTATATCTGATGAGCTG 2910
 DB 2640 TGACCATGATGCTCAAGAGAGTGTGATATATCCATGAGTATATCTGATGAGCTG 2699
 QY 2911 ACTGGTTATTAAGAGACTACCCACATACATGATGATGATGATGATGATGATGATG 2970
 DB 2700 ACTGGTTATTAAGAGACTACCCACATACATGATGATGATGATGATGATGATGATG 2759
 QY 2971 GTCTTGGGGTTGGGTAGGTAGATACCAATGAACACCTTTCAGAGCCCTTCTCTCT 3030
 DB 2760 GTCTTGGGGTTGGGTAGGTAGATACCAATGAACACCTTTCAGAGCCCTTCTCTCT 2819
 QY 3031 GCAGTGTCTTCTTAACTCTCTTACTAGAGAGATATATTTTGCATATATGAAGAA 3090
 DB 2820 GCAGTGTCTTCTTAACTCTCTTACTAGAGAGATATATTTTGCATATATGAAGAA 2879
 QY 3091 TTTTCTAGTATATTAACGAGCCCTTATTTCTAAATGATGATGATGATGATGATG 3150
 DB 2880 TTTTCTAGTATATTAACGAGCCCTTATTTCTAAATGATGATGATGATGATGATG 2939
 QY 3151 AGGATTAACAGATGATTTTGAATTTCCAGAGATATTAATGATGATGATGATG 3210
 DB 2940 AGGATTAACAGATGATTTTGAATTTCCAGAGATATTAATGATGATGATGATG 2999
 QY 3211 ATAAATCATCTTGTCT 3227
 DB 3000 ATAAATCATCTTGTCT 3016

RESULT 3
 ID ABA83077 standard; DNA; 2598 BP.
 XX ABA83077;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE Human transcription factor TRFX-104 coding sequence.
 XX
 KW Human; transcription factor; TRFX; cell proliferative disease;
 KW autoimmune disease; inflammation; neurological disease;
 KW developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;
 KW neuroprotective; antiinflammatory; gene therapy; ds.
 XX
 OS Homo sapiens.
 PN W0200172777-AZ.
 XX
 PD 04-OCT-2001.
 XX
 PF 13-MAR-2001; 2001WO-US08117.
 XX
 PR 13-MAR-2000; 2000US-0188986.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 PI Hillman JI, Baughn MR, Yue H, Lai P, Lu DW, Patterson C;
 PI Azimzal Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;

QY 2436 ATACTTCACCAAGACCTCAAAACCTGGCAGAGAAGGCCAAGAAATCTGTCGACA 2495
 |||||||
 DB 1792 ATACTTCACCAAGACCTCAAAACCTGGCAGAGAAGGCCAAGAAATCTGTCGACA 1851
 QY 2496 TAAGCTGATCTATGTTGAGGCTGGAAGTACTCACAGGCTTCTTCTGCGAGGACATCC 2555
 |||||||
 DB 1852 TAAGCTGATCTATGTTGAGGCTGGAAGTACTCACAGGCTTCTTCTGCGAGGACATCC 1911
 QY 2556 GTTAATATGTTGTTCCCAACCTGAGTCCCTGAGCTATAGATGTCATGAGAGTGTGT 2615
 |||||||
 DB 1912 GTTAATATGTTGTTCCCAACCTGAGTCCCTGAGCTATAGATGTCATGAGAGTGTGT 1971
 QY 2616 TGACATGCTCCAGCCCTGGAACCTTGTATTTTATGAGCATGACAGCTATACCTCA 2675
 |||||||
 DB 1972 TGACATGCTCCAGCCCTGGAACCTTGTATTTTATGAGCATGACAGCTATACCTCA 2031
 QY 2676 GTATTTGTGACATGACCCAGGAATATGTACAGTACACAGATTAATCTATCAGGACCAAT 2735
 |||||||
 DB 2032 GTATTTGTGACATGACCCAGGAATATGTACAGTACACAGATTAATCTATCAGGACCAAT 2091
 QY 2736 GAGACCTTTGACAGCATCCTTGGAGGGGATTCGCAAGTGGCAGCCCTGTGATTATC 2795
 |||||||
 DB 2092 GAGACCTTTGACAGCATCCTTGGAGGGGATTCGCAAGTGGCAGCCCTGTGATTATC 2151
 QY 2796 TCTACATCATTTTAAACAGCTGATGCGCTTACCTTGGGTGAACCAATTAATGACC 2855
 |||||||
 DB 2152 TCTACATCATTTTAAACAGCTGATGCGCTTACCTTGGGTGAACCAATTAATGACC 2211
 QY 2856 ATCGATGCGCTCAAGAGTGGCTGAATATATCCATGGGTTATCTGTATGACTGACTGG 2915
 |||||||
 DB 2212 ATCGATGCGCTCAAGAGTGGCTGAATATATCCATGGGTTATCTGTATGACTGACTGG 2271
 QY 2916 GTTATTTGAAGAGACTAGCCACATAGTCTTGTAGTGTCTTATCTGCTTATGTCCT 2975
 |||||||
 DB 2272 GTTATTTGAAGAGACTAGCCACATAGTCTTGTAGTGTCTTATCTGCTTATGTCCT 2331
 QY 2976 GGGGTTGGGGTAGTAGATACCAATGAACACTTTGACAGGACCTTCCCTCTTGACAGT 3035
 |||||||
 DB 2332 GGGGTTGGGGTAGTAGATACCAATGAACACTTTGACAGGACCTTCCCTCTTGACAGT 2391
 QY 3036 TCTCTTTATCTCCTTTACTAGAGAGATTAATATTTTGCATATATAGTAAAGAAATTTT 3095
 |||||||
 DB 2392 TCTCTTTATCTCCTTTACTAGAGAGATTAATATTTTGCATATATAGTAAAGAAATTTT 2451
 QY 3096 CTAGATATATAGCGAGCGCTTTATTTCTAAATGATATATATTAATAAAGTTAGGAT 3155
 |||||||
 DB 2452 CTAGATATATAGCGAGCGCTTTATTTCTAAATGATATATTAATAAAGTTAGGAT 2511
 QY 3156 AACAGATGATTTTGAATTTCCAGAGATTTATTAAGTCTTATAGTAAAGTAAATAA 3215
 |||||||
 DB 2512 AACAGATGATTTTGAATTTCCAGAGATTTATTAAGTCTTATAGTAAAGTAAATAA 2571
 QY 3216 TCATCTTTGCTGATTAATAAAAAAAA 3242
 |||||||
 DB 2572 TCATCTTTGCTGATTAATAAAAAAAA 2598
 |||||||
 RESULT 4
 AB054419 standard; cDNA; 2581 BP.
 ID AB054419;
 AC AB054419;
 XX 22-AUG-2002 (first entry)
 DT 22-AUG-2002 (first entry)
 XX Human ovarian antigen HTXL31 cDNA, SEQ ID NO:299.
 DE Human ovarian antigen HTXL31 cDNA, SEQ ID NO:299.
 XX Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;

KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; gene; ss.
 OS Homo sapiens.
 PN W0200200677-A1.
 PD 03-JAN-2002.
 PF 07-JUN-2001; 2001WO-US18569.
 PR 07-JUN-2000; 2000US-209467P.
 PA (HUMA-) HUMAN GENOME SCT INC.
 PI Birse CE, Rosen CA;
 XX WPI: 2002-147878/19.
 DR P-PSDB; ABP41342.
 DR Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 PS Claim 1; SEQ ID No 299; 2922pp; English.
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (AB054131-AB056305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosis or preventing various cancer and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 XX
 SQ Sequence 2581 BP; 843 A; 478 C; 574 G; 685 T; 1 other:
 Query Match 78.1%; Score 2533; DB 24; Length 2581;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2544; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 QY 697 GGTGATTTGAATCCAAAGAGAGCAAGCTTGTGCGAGATATGCTAAAGTGCA 756
 |||||||
 DB 1 GGTGATTTGAATCCAAAGAGAGCAAGCTTGTGCGAGATATGCTAAAGTGCA 60
 |||||||
 QY 757 GCTGTGAGATAGCTGTCACGGAGACAGGAGGCTTCCCTGCAACAGATCATCCATGCT 816
 |||||||

D 61 GCTGTGATAGCTGTACAGGAGCAGGAGGCTTCCTGCAAAACGATCATCATCT 120
Q 817 GTTGGGCTGTGGATGGAATGAACAGGAGATGTACTGGAAAGCTGAGAGGGCC 876
D 121 GTTGGGCTGTGGATGGAATGAACAGGAGATGTACTGGAAAGCTGAGAGGGCC 180
Q 877 ATTGAATATTCGATTAATGATCTATATAAAATCTACATTAAGACAGTACAT 936
D 181 ATTGAATATTCGATTAATGATCTATATAAAATCTACATTAAGACAGTACAT 240
Q 937 CCAGCCTTGAGCTGTGGATTTTTCAGTTCCTCTGTAATTTGTACAAAGACTATTG 996
D 241 CCAGCCTTGAGCTGTGGATTTTTCAGTTCCTCTGTAATTTGTACAAAGACTATTG 300
Q 997 GAGACTATCGGGTTAGTTTGAAGGAGGAGCAATGATGATTAATTTGAAGAATTCAC 1056
D 301 GAGACTATCGGGTTAGTTTGAAGGAGGAGCAATGATGATTAATTTGAAGAATTCAC 360
Q 1057 CTGGTGAGCAATGAGACCCCTACTGTGCTGCTTTAAAGCTGCTTCAAGATTCATCT 1116
D 361 CTGGTGAGCAATGAGACCCCTACTGTGCTGCTTTAAAGCTGCTTCAAGATTCATCT 420
Q 1117 GGGAAAGTAGCTGGGACAAAGAACCCCTTCTTCAATGCAATGTCGTGAACAC 1176
D 421 GGGAAAGTAGCTGGGACAAAGAACCCCTTCTTCAATGCAATGTCGTGAACAC 480
Q 1177 CTGACCTTCAGATTTCTCAGGCGCACATTTGAATGCGACGCGAGATGTAATTT 1236
D 481 CTGACCTTCAGATTTCTCAGGCGCACATTTGAATGCGACGCGAGATGTAATTT 540
Q 1237 TCTGTAAACCCACATGATTAATTAAGTTGACCTGTGCAAAAGTCATTAATTCAC 1296
D 541 TCTGTAAACCCACATGATTAATTAAGTTGACCTGTGCAAAAGTCATTAATTCAC 600
-Q 1297 GGGTTGAATGAATGGGAATTTCTTCCACAAAGGCTAAACGTTTCAACGCTCCAC 1356
D 601 GGGTTGAATGAATGGGAATTTCTTCCACAAAGGCTAAACGTTTCAACGCTCCAC 659
Q 1357 TTGGTACTGTGCAAAAGATTTAACTTGTCTGTAATATATATACATGTACTGTG 1416
D 660 TTGGTACTGTGCAAAAGATTTAACTTGTCTGTAATATATATACATGTACTGTG 719
Q 1417 CATTCAGATTTCTTAAACCTCAGATATTAACATGCAATGAAGAGTGTGGAAAA 1476
D 720 CATTCAGATTTCTTAAACCTCAGATATTAACATGCAATGAAGAGTGTGGAAAA 779
Q 1477 TGCATTGAGAAAAATTAATCTCATTTCTTCTTCCCTTGGGAGTGAACATGGAA 1536
D 780 TGCATTGAGAAAAATTAATCTCATTTCTTCTTCCCTTGGGAGTGAACATGGAA 839
Q 1537 ATAAAGAGAAACAGCAGCAGAGATTTTGTGATGAAGTTTAAACATTTGCCAAAGAC 1596
D 840 ATAAAGAGAAACAGCAGCAGAGATTTTGTGATGAAGTTTAAACATTTGCCAAAGAC 899
Q 1597 CATGTAAACACAGTTAACTGTAAATTTGTGATCTTTCCACAGATTTTGGAGATAT 1656
D 900 CATGTAAACACAGTTAACTGTAAATTTGTGATCTTTCCACAGATTTTGGAGATAT 959
Q 1657 AAGGCTTTCAGTTCTGAAATGGCAAGAGTCCAGATGCTGAGTTTGAACATTAAGT 1716
D 960 AAGGCTTTCAGTTCTGAAATGGCAAGAGTCCAGATGCTGAGTTTGAACATTAAGT 1019
Q 1717 GTCCCCAGTCAACAGAGAGAGAAAAAGAAATGGCTTGAAGCTAGATCTCTGCC 1776
D 1020 GTCCCCAGTCAACAGAGAGAGAAAAAGAAATGGCTTGAAGCTAGATCTCTGCC 1079
Q 1777 ATCAATCTGATGGATTCACAGTGAAGAGATGTATGAGGCCACGATGATCCAAAGA 1836
D 1080 ATCAATCTGATGGATTCACAGTGAAGAGATGTATGAGGCCACGATGATCCAAAGA 1139
Q 1837 ATCTGTGATCTCCAGAACCAACACATATGAGAAATATATATCTGTACTTTGGGAGA 1896
D 1140 ATCTGTGATCTCCAGAACCAACACATATGAGAAATATATATCTGTACTTTGGGAGA 1199

Q 1897 AAGAACATGACATTTTGTCTCAGCTTCAGAAAACTCAAGTGTCTCATCAGAAAT 1956
D 1200 AAGAACATGACATTTTGTCTCAGCTTCAGAAAACTCAAGTGTCTCATCAGAAAT 1259
Q 1957 ATCAGCCAGAGAGACAGAGATTAGAGATTGAAGAGACCCGGGCTGACCTATTGAGCTG 2016
D 1260 ATCAGCCAGAGAGACAGAGATTAGAGATTGAAGAGACCCGGGCTGACCTATTGAGCTG 1319
Q 2017 GTTATGACATTTGAAGTATGCTTTTGAAGTACAGAGGAATGCGAAGGAAAAAGAG 2076
D 1320 GTTATGACATTTGAAGTATGCTTTTGAAGTACAGAGGAATGCGAAGGAAAAAGAG 1379
Q 2077 CGAGCCCTTGGCGCTGTTAGACAGTGAACATTTAGCAACAAAAACCCAAAGACGA 2136
D 1380 CGAGCCCTTGGCGCTGTTAGAGACAGTGAACATTTAGCAACAAAAACCCAAAGACGA 1439
Q 2137 ATGAAGAAATATCATATTTTGAAGTGTCTGTGCTTCAACTCAAGAGCTTCTAGAT 2196
D 1440 ATGAAGAAATATCATATTTTGAAGTGTCTGTGCTTCAACTCAAGAGCTTCTAGAT 1499
Q 2197 CAAAAGAAACAGTTTGAAGAAATGCTTGCAGTTCTAAAGGTTGGAAGATAGACAT 2256
D 1500 CAAAAGAAACAGTTTGAAGAAATGCTTGCAGTTCTAAAGGTTGGAAGATAGACAT 1359
Q 2257 GAGTCTCTATGCTGCTGCTTCAAGAAAGAAAGAAATGAGAGAAAACTGCACAGG 2316
D 1560 GAGTCTCTATGCTGCTGCTTCAAGAAAGAAAGAAATGAGAGAAAACTGCACAGG 1619
Q 2317 CAACCTGTGAGCCATAGGCTGTTTCAAGAGTCCCATACAGTTTCTCAATGTGTATGC 2376
D 1620 CAACCTGTGAGCCATAGGCTGTTTCAAGAGTCCCATACAGTTTCTCAATGTGTATGC 1679
Q 2377 AGAGTTGGCTTCAAGAAATGATCTGACACCTTGGCATCAAAATPGAGAGTGGCAT 2436
D 1680 AGAGTTGGCTTCAAGAAATGATCTGACACCTTGGCATCAAAATPGAGAGTGGCAT 1739
Q 2437 TACTTACCAAGAACCTCAAAAACTTGCAGAGAGAGCCAAAGAAATCTGTCTGACAT 2496
D 1740 TACTTACCAAGAACCTCAAAAACTTGCAGAGAGAGCCAAAGAAATCTGTCTGACAT 1799
Q 2497 AAGCTGATCTATGTTTGAAGCTGAAGTACTCAGAGGCTTCTTCCACAGGACATCCG 2556
D 1800 AAGCTGATCTATGTTTGAAGCTGAAGTACTCAGAGGCTTCTTCCACAGGACATCCG 1859
Q 2557 TTAATATTTGTCCTCCACACAGTGAAGTCTGAGGCTATGATGGTCAATGAGTGGT 2616
D 1860 TTAATATTTGTCCTCCACACAGTGAAGTCTGAGGCTATGATGGTCAATGAGTGGT 1919
Q 2617 GACATGTCCTCAGACCTGGAACCTTGTATTTTATGTCATGACAGGCTATACCTCAG 2676
D 1920 GACATGTCCTCAGACCTGGAACCTTGTATTTTATGTCATGACAGGCTATACCTCAG 1979
Q 2677 TATTTTGAACATGACCCAGAAATATGTACAGTACACATTAATCATAGACCAATG 2736
D 1980 TATTTTGAACATGACCCAGAAATATGTACAGTACACATTAATCATAGACCAATG 2039
Q 2737 AGACCCCTTGCACAGCATCTTGGAGGAGATTTGCAAGTGGACGCGCTGTTGATTAATCT 2796
D 2040 AGACCCCTTGCACAGCATCTTGGAGGAGATTTGCAAGTGGACGCGCTGTTGATTAATCT 2099
Q 2797 CTACATCATTTTAAACAGCTGTATGCTTACCTTGGGTAACCTAACCAATATAGCA 2856
D 2100 CTACATCATTTTAAACAGCTGTATGCTTACCTTGGGTAACCTAACCAATATAGCA 2159
Q 2857 TCGATGCTTAAAGAGTGGCTTGAATATATCCATGAGGCTTATCTGTATGACATGAC 2916
D 2160 TCGATGCTTAAAGAGTGGCTTGAATATATCCATGAGGCTTATCTGTATGACATGAC 2219
Q 2917 TTAATGAAAGACCTAGCAACATCTAGACATCTTATGCTTATCTTATGCTTATGCTG 2976
D 2220 TTAATGAAAGACCTAGCAACATCTAGACATCTTATGCTTATCTTATGCTTATGCTG 2279

Db 691 TCCGTTAAATATTTGTTCCCCACACTGAGTCTCTGAGCTATAGATGTCATGACAGTGT 632
QY 2613 GATTGACAAATGTCCTCCAGCCCTGAACCTTTGTTATTTTGTAGTCAGCAGCTATACC 2672
Db 631 GGTGGCAATGTCCTCCAGCCCTGAACCTTTGTTATTTTGTAGTCAGCAGCTATACC 572
QY 2673 TCAGTATTTGTGACATGCAACCCAGAAATATGACAGTCAACAAGATTACTCATCAGACC 2732
Db 571 TCAGTATTTGTGACATGCAACCCAGAAATATGACAGTCAACAAGATTACTCATCAGACC 512
QY 2733 AATGAGACCCCTTTGACAGATCCTTGGAGGGATTCGCAAGGGAGAGCCCTTGATTA 2792
Db 511 AATGAGACCCCTTTGACAGATCCTTGGAGGGATTCGCAAGGGAGAGCCCTTGATTA 452
QY 2793 ATCTACATCATTTTAAACAGCTGTATGGCCTTACCTTGGGTGAGACTAACAATAATAG 2852
Db 451 ATCTACATCATTTTAAACAGCTGTATGGCCTTACCTTGGGTGAGACTAACAATAATAG 392
QY 2853 ACCATGATGCTCAAAAGAGTGGCTTGAATATATCCCATGGTTATCTGTATGAGCTGAC 2912
Db 391 ACCATGATGCTCAAAAGAGTGGCTTGAATATATCCCATGGTTATCTGTATGAGCTGAC 332
QY 2913 TGGGTATTTGAAGGACTGACCATACATACATCTAGGCTTATCTGCTTATGCTTATGT 2972
Db 331 TGGGTATTTGAAGGACTGACCATACATACATCTAGGCTTATCTGCTTATGCTTATGT 272
QY 2973 CTGGGGTGGGGTAGTAGATACCAATGAACACTTTCAGAGCCCTCCCTCTCTGC 3032
Db 271 CTGGGGTGGGGTAGTAGATACCAATGAACACTTTCAGAGCCCTCCCTCTCTGC 212
QY 3033 AGTTGTTCTTTAAATCTCTTACTAGAGAGATAAATTTTGCATATATTAAGAATTT 3092
Db 211 AGTTGTTCTTTAAATCTCTTACTAGAGAGATAAATTTTGCATATATTAAGAATTT 152
-QY 3093 TTCTCTGATATATACGAGGCTTTATTTCTTAAATGATGATAGATTAAGATTTAG 3152
Db 151 TTCTCTGATATATACGAGGCTTTATTTCTTAAATGATGATAGATTAAGATTTAG 92
QY 3153 GATAACAGATGATTTTGTAGATTTTCCAGAGAAATATTATAAGTCTTTAGTATGAAT 3212
Db 91 GATAACAGATGATTTTGTAGATTTTCCAGAGAAATATTATAAGTCTTTAGTATGAAT 32
QY 3213 AATTCATCTTTGCTGATTAATAAAAAA 3243
Db 31 AATTCATCTTTGCTGATTAATAAAAAA 1
RESULT 6
AAK51529
ID AAK51529 standard; cDNA; 1747 BP.
XX
AC AAK51529;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 74.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001MO-US04098.
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
DR WPI: 2001-476283/51.
DR P-PSDB: AAM78396.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -
XX
PS Claim 1: Page 675-677; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAH80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation on which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAH80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 1747 BP; 546 A; 342 C; 399 G; 460 T; 0 other:
Query Match 47.1%; Score 1526.2; DB 22; Length 1747;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1528; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1713 CAGTGTCCCCCACTCAACCCAGAGAGAGAAAGAAATGGGCTTGAACTGATCTCC 1772
Db 217 CAAAGTCCCCCACTCAACCCAGAGAGAGAAAGAAATGGGCTTGAACTGATCTCC 276
QY 1773 TGCCATCAATCTGATGATGATTCACCTGGAAGAGATGTAAGAGCCACGATGATCCA 1832
Db 277 TGCCATCAATCTGATGATGATTCACCTGGAAGAGATGTAAGAGCCACGATGATCCA 336
QY 1833 AAGAATCTGAGTCTCCAGAACCCACACATCATTTGAATATCATATTTCTGACCTGG 1892
Db 337 AAGAATCTGAGTCTCCAGAACCCACACATCATTTGAATATCATATTTCTGACCTGG 396
QY 1893 GAGAAAGAACATGACATTTTGTCTCAGCTTCAGAAACCTTCAAGTGTCTCCATCAGAG 1952
Db 397 GAGAAAGAACATGACATTTTGTCTCAGCTTCAGAAACCTTCAAGTGTCTCCATCAGAG 456
QY 1953 AATTATCAGCCCGAGAGAGCAGAGTTAGAGATTGAAGAGCCCGGGCTGACCTCTTGA 2012
Db 457 AATTATCAGCCCGAGAGAGCAGAGTTAGAGATTGAAGAGCCCGGGCTGACCTCTTGA 516
QY 2013 GGTGTTATGAAACATGAGATATGCTTTGTAAGTACAGAGAGAAATGGCAAGAAAAA 2072
Db 517 GGTGTTATGAAACATGAGATATGCTTTGTAAGTACAGAGAGAAATGGCAAGAAAAA 576
QY 2073 GAGAGAGGCTTTGGCGCTGTTAGACAGTGAATTCAGCAACAACCAAGAG 2132
Db 577 GAGAGAGGCTTTGGCGCTGTTAGACAGTGAATTCAGCAACAACCAAGAG 636
QY 2133 CGAAATGAAGAAATATCATATTTCTGAAGATGCTGTCGCTCAACATCAAGAGCTTC 2192
Db 637 CGAAATGAAGAAATATCATATTTCTGAAGATGCTGTCGCTCAACATCAAGAGCTTC 696

QY 2193 AGATCAAAAGAACAGTTTGAAATATGTGTTGCGAGTCTCTAAAGGTGAGAAAGATAGA 2252
 |||||||
 Db 697 AGATCAAAAGAACAGTTTGAAATATGTGTTGCGAGTCTCTAAAGGTGAGAAAGATAGA 756
 QY 2253 CAATGAGGTCTTATGCTGCTTCAAGAAAGAAATGATGAGAAATCACTCA 2312
 |||||||
 Db 757 CAATGAGGTCTTATGCTGCTTCAAGAAAGAAATGATGAGAAATCACTCA 816
 QY 2313 CAGCAACCTGTAGCCATAGGCTGTTCAGCAAGTCCCATACAGTCTCTCATATGTGT 2372
 |||||||
 Db 817 CAGCAACCTGTAGCCATAGGCTGTTCAGCAAGTCCCATACAGTCTCTCATATGTGT 876
 QY 2373 ATGCAAGATTGCTTCAAGAAATGTACTCAGACCTTCCGCTCAAAATACGAGCTGG 2432
 |||||||
 Db 877 ATGCAAGATTGCTTCAAGAAATGTACTCAGACCTTCCGCTCAAAATACGAGCTGG 936
 QY 2433 CATATACTTCACCAAGACCTCAAAACCTGGCAGAGAGCCCAAGAAATCTCTGCTGC 2492
 |||||||
 Db 937 CATATACTTCACCAAGACCTCAAAACCTGGCAGAGAGCCCAAGAAATCTCTGCTGC 996
 QY 2493 AGATTAAGTGTATCTGTGTGTTGAGGCTGAAAGTACTCAGAGCTTCTTCTGCGAGGACA 2552
 |||||||
 Db 997 AGATTAAGTGTATCTGTGTGTTGAGGCTGAAAGTACTCAGAGCTTCTTCTGCGAGGACA 1056
 QY 2553 TCCGTTAAATATGTGTTCCCGCACCACTGAGTCCCTGAGCTATAGATGGTCATGACAGTGT 2612
 |||||||
 Db 1057 TCCGTTAAATATGTGTTCCCGCACCACTGAGTCCCTGAGCTATAGATGGTCATGACAGTGT 1116
 QY 2613 GGTGTACAAATGTCTCCAGCCCTGAAACCTTGTATTTTATAGTGGCATGACAGCTATAC 2672
 |||||||
 Db 1117 GGTGTACAAATGTCTCCAGCCCTGAAACCTTGTATTTTATAGTGGCATGACAGCTATAC 1176
 QY 2673 TCAGTATTTGGGACATCAGTCCAGGAAATGTATGTACAGTCAAGATTTACTCATCGAGACC 2732
 |||||||
 Db 1177 TCAGTATTTGGGACATCAGTCCAGGAAATGTATGTACAGTCAAGATTTACTCATCGAGACC 1236
 QY 2733 AATGAGACCTTGTGACAGCATCTCTGGAGGGATTTCCAGTGGCAGCCCTGTGATTA 2792
 |||||||
 Db 1237 AATGAGACCTTGTGACAGCATCTCTGGAGGGATTTCCAGTGGCAGCCCTGTGATTA 1296
 QY 2793 ATCTCTACATCATTTTAAACAGCTGTATGCTTACCTTGGGTGAATACCAAAATATG 2852
 |||||||
 Db 1297 ATCTCTACATCATTTTAAACAGCTGTATGCTTACCTTGGGTGAATACCAAAATATG 1356
 QY 2853 ACCATCGATGGCTCAAAAGAGTGGCTGTGAATATATCCCATGGTATCTGTATGACTGAC 2912
 |||||||
 Db 1357 ACCATCGATGGCTCAAAAGAGTGGCTGTGAATATATCCCATGGTATCTGTATGACTGAC 1416
 QY 2913 TGGGTATTTGAAAGGACTAGGACACATCTAGCATCTTATGCTTATCTGTATATGT 2972
 |||||||
 Db 1417 TGGGTATTTGAAAGGACTAGGACACATCTAGCATCTTATGCTTATCTGTATATGT 1476
 QY 2973 CTTGGGGTTGGGGTAGTAGATACCAAAATGAACACTTTCAGGACCTTCTCTCTTGC 3032
 |||||||
 Db 1477 CTTGGGGTTGGGGTAGTAGATACCAAAATGAACACTTTCAGGACCTTCTCTCTTGC 1536
 QY 3033 AGTGTCTTTAAATCTCTTAACTAGAGAGATATAATATTTTGCATATATGAAATAT 3092
 |||||||
 Db 1537 AGTGTCTTTAAATCTCTTAACTAGAGAGATATAATATTTTGCATATATGAAATAT 1596
 QY 3093 TTTCTAGATATTAACGAGGCTTTATTTCTTAAATGATCATGTATTAATATGTTAG 3152
 |||||||
 Db 1597 TTTCTAGATATTAACGAGGCTTTATTTCTTAAATGATCATGTATTAATATGTTAG 1656
 QY 3153 GATAACAGAAATGATTTAGATTTTCCAGAGAAATATTAAGTGTAGGATGAATAAT 3212
 |||||||
 Db 1657 GATAACAGAAATGATTTAGATTTTCCAGAGAAATATTAAGTGTAGGATGAATAAT 1716
 QY 3213 AAATCATCTTGTCTGATTAATAAAAAAAAAA 3243
 |||||||
 Db 1717 AAATCATCTTGTCTGATTAATAAAAAAAAAA 1747

RESULT 7
 AAD00630
 ID AAD00630 standard; cDNA: 3024 BP.
 XX
 AC AAD00630;
 XX
 DT 29-AUG-2000 (first entry)
 DE Mouse B-aggressive lymphoma (BAL) protein-short form cDNA.
 KW Mouse; B-aggressive lymphoma; BAL; tumour; malignancy;
 KW differential expression; DLB-CL; Diffuse large B-cell lymphoma;
 KW cytosolic; vaccine; treatment; diagnosis; non-Hodgkin's lymphoma;
 KW cellular adhesion; sarcoma; carcinoma; myeloma; ss.
 XX
 OS Mus sp.
 FH
 FH Key Location/Qualifiers
 FT 5'UTR 1..170
 FT CDS /*tag= a
 FT 171..2651
 FT /*tag= b
 FT /product= "BAL protein-short form"
 FT 171..2648
 FT /*tag= c
 FT /note= "This region is specifically claimed in Claim 1
 as SEQ ID NO:6"
 FT 3'UTR 2652..3024
 FT /*tag= c
 FT misc_feature 2295..2739
 FT /*tag= d
 FT /note= "This region is 99% identical to Soares 2NBMF
 Mus musculus cDNA clone 1446050"
 FT
 PN W0200026231-A1.
 XX
 PD 11-MAY-2000.
 XX
 PE 29-OCT-1999; 99WO-US25439.
 XX
 PR 29-OCT-1998; 98US-0106383.
 PR 30-OCT-1998; 98US-0106448.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 PI Shipp M, Aguilar R, Yakushijin Y;
 XX
 DR WPI: 2000-365563/31.
 DR P-PSDB: AAY71083.
 PT New B-aggressive lymphoma nucleic acid for identifying cells exhibiting
 PT or predisposed to malignancies such as lymphoma, sarcoma, carcinoma and
 PT myeloma -
 PS
 PS Claim 1; Fig 2; 151pp; English.
 XX
 CC The present sequence is a cDNA encoding short form of B-aggressive
 CC lymphoma (BAL) protein obtained from mouse (Balb-c) spleen by 5' and 3'
 CC RACE using a 418 bp clone as an anchor from Soares mouse mammary gland
 CC having homology to human sequence and primers from mouse EST database.
 CC BAL was identified to be differentially expressed in DLB-CL (Diffuse
 CC large B-cell lymphoma), the most common non-Hodgkin's lymphoma, by
 CC differential display technique. It was found to be more abundant in
 CC tumours from patients with high risk fatal DLB-CL disease than low risk
 CC cured disease. BAL transcripts are most abundant in lymphoid organs e.g.
 CC spleen, foetal liver and peripheral blood, and several non-haematopoietic
 CC organs e.g. heart and skeletal muscle. BAL is involved in modulation of
 CC cellular adhesion and aggressiveness/severity of malignancy such as
 CC DLB-CL. BAL nucleic acid and protein are useful for identifying cells
 CC exhibiting or predisposed to malignancies such as lymphoma, sarcoma,
 CC carcinoma and myeloma. BAL proteins are useful as immunogens to raise
 CC anti-BAL antibodies. BAL modulators or BAL molecules can be used for
 CC prophylactic and therapeutic treatment of a subject susceptible to or

CC having a disorder associated with aberrant BAL expression or activity,
CC such as non-Hodgkin's lymphoma. It is also useful in vaccine preparation.
XX
Sequence 3024 BP; 902 A; 635 C; 725 G; 762 T; 0 other;

Query Match 41.3%; Score 1340; DB 21; Length 3024;
Best Local Similarity 70.5%; Pred. No. 0;
Matches 2053; Conservative 0; Mismatches 735; Indels 122; Gaps 15;

```

367 CCTACAGAGCTCAGAGAGCTGCTCTGAGAAAAACATAGTGGCAATTCCTCAT 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
192 CCGCGGAAAGACAGACACATCTCTGAGAACATTATAGATGGCAATTCCTCAT 251
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
427 AACCAAGTACTTCAAAATTTTAAAAATATAGAGCTGCTGAGAGCTCCAC 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
252 AACCAAGTACTTCAAAATTTTAAAAAGAGCAATGAGAGCTGCTGAGAGCTCC 311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
487 AATAAGTTGGCTGATCTCTACCTGCTGCTCCAGTTCCAGAGGACACAGC---AA 543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
312 AATAGTTGGATGATCTCTACCTGAGCTGATCCAACTTACAGAGGACACCTCTCT 371
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
544 TCTGTCAAGTGTTCAGAAAAATGCTGACCTAGATAGATTATCAGCTGAGAAAG 603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
372 GCTCAGAGAGTCTTCAGAGAGAGAGCTGATCCCTGGGATGAGTTATCTGTGGA 431
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
604 GACCTACACACATGCTGTGATGCTGTGATGATGCAATGCAATGAGATCTTGCAT 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
432 GACCTTACACAGACAGCTGTGATGCTGTGATGCAACGACCAATGAAACCTTT 491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
664 GGGGAGAGCTGGGCTGGGCTGGTGAAGCTGGTGGATTTGAAATCCAGAGAGAG 723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
492 GAGAGAGCTGGGCTGGGCTGGTGAAGCTGGTGGATTTGAAATCCAGAGAGAG 551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
724 AACAGAGTTGCTCCAGATATGTTAAAGTGCAGCTGCTGATAGCTGACAGAGAG 783
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
552 AAAAGATCTATTCAGACGTTGTAATATCTGATGGTGAATCGCTATCACCGGTG 611
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
784 GGGAGGCTCCCTGCAAAAGATCATCATGCTGTGGGCTGGTGGATGAGAGAGAT 843
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
612 GGGAGAGCTCTGCTGCAATTTGATATCCATGCGGTTGAGACCTGGTGGACAG 671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
844 AACAGAGATGATGATGGAAGCTGACAGAGGCTGATGATGATGATGATGATGAT 903
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
672 AGCAGAGACGCTATCGAATTAAGAAATTTGCAATTAAGAACTTATGTCAC 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
904 TATATAAATCTCAGATTAAGACAGTACGATTCAGACCTGAGCTGAGATTTTTC 963
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
732 AATATGATCTACGATTAAGACAGTACGATTCAGACCTGAGCTGAGATTTTTC 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
964 TTTCTCTGATTTGTGTACAAAGACTATTTGTAGAGACTATCCGGGTTTACTT 1023
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
792 TTTCTCTGATTTGTGTACAAAGCATTAATTTTGAAGAACTATTCGCGGTTT 851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1024 AACCAATGATGATTAATTTGAAGAAATTCACCTGAGTACGATGAGACCTACT 1083
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
852 AACCAATGATGATTAATTTGAAGAAATTCATCTGTGAGCAATGAGACCTACT 911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1084 GTCGCTTTAAAGCTGCTGAGATTCATCTGAGGAGAGAG-----TGAGCT 1134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
912 GCTCTCTTAATATCCGCTGAGAAAGCATCTTGGAGGAGAGACTGAGCTTTT 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1135 CAGAAACACACCCCTCTT-----TCATGCAATGATGCTGAAACACCTGAGCT 1188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
972 CAGAAACACACCCCTCTTCCACATGATGATCTTGCATGCGGCGGCGCTGAG 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1189 ATTGTCCAGGCGCACATTTGATGAGAGGCGAGATGATTAATTTCTGTAACCCA 1248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1032 ATTGTCCAGGCGCTGATTTGAAGAAACCAACAGATGATTAATTTCTGATCAT 1091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1249 CAGTATATTACGATTTGAGAGCTGAGCAAGTCAATCTACACAGAGAGAGTTGA 1308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1092 CAGGATTTTAATCAGACAGAGTGGCACAGTCAATCTTACAGACAGAGGAGTT 1151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

1309 AATTCGATTTCTTGCCACAAAGCTAAACGATTTCAACGGTCCAGTTGGTACTG 1368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1152 GAAAGGAACTT-----GACAAGGTTAACTCTCCACAGATTAATCAAGAGTGG 1205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1369 ACAAAAGATTTAATGTTCTGTAAATATATATACCATGTACTGCGCATTCAGAA 1428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1206 ACAAAAGATTTAATGTTCTGTCAATGTTCTTCCATGTCGATGAGATTCCTCAA 1265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1429 CTTAACTCTGATTAATTTAAACATGCAATGCAAGAGAGTGTGAAAAATTCATT 1488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1266 AACAAATACAGATTAATGAAAGATGCAATGAAATGCTGTACGAAATAATCTTAA 1325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1489 AATATTAATCTTCCATTTCTTCTGCTGCGCTGGGACGAGCAATGGAATTAAGA 1548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1326 GATTAATTAATCTTCCATTTCTTCTGCTGCGGACAGAGATGATGGAATTTGA 1385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1549 ACAGCAGAGAGATTTGTTGATGAGTTTAACTTTGCTCCAAAGACCTGTAACA 1608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1386 ACAGCAGCTCAGATTAATGTTTGAAGAGATTTTGCATTTGCTTAAAGACCA 1445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1609 CAGTTAACTGTAAATTTTGTGATCTTCCAAACAGATTTGAGATTAATAGCTT 1668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1446 AGCTTAACGTGTAAGATTTGATCTTTCCAGTATGATGAGACGTAACAAGATT 1505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1669 TCTGAATGSCAAAGAGTCCAAAGATGCTGATGATTTGAACATTTACAGT----- 1719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1506 GCTGAATATGCAAAAGAGTCCAAAGAGCTGAATCTAGCGCTAATATGCTGCT 1565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1720 CCCCAGTCAACGAGAGAGGAAAGAAATGGGCTTGAACCTGATGATCTGCTG 1779
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1566 CTGACAGTGTCCAGTGGGAGGCAAGAGAGGCGGCTTGAACGATGATCTGCTG 1825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1780 AATCTATGATTTCAACGCTGAGAGAGATGATGAGCGCCACCATGATCCAAAGA 1839
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1626 AATCTATGATTTCAAGTGGAGAGATGATGAGAGCCACGAGATGATTAAGATT 1885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1840 CTGAGTCTCCAGAACACACATCATTTGATTAATCATTTTGTACCTTGGGAGAG 1899
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1686 CTGCTCTCCCTGAGACACCATCATTTGATTAATCATTTTGTACCTTGGGAGAG 1745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1900 GATCATGACATTTTGTCTGCTGCAAGAACTTCAAGTGTCTCCATCATCAGAA 1959
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1746 GAGCAGAGAGCTGTGTGCTGCTGCAAGACGACAGCAAGAGATCTCATTTGAG 1805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1960 AGCCAGAGAGAGACAGATTTAGAGATTTGAAGAGCCGCGGCTGACCTATTTG 2019
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1806 AGTCCAGAAAGGCGCACTTTGAGATTTAAAGTCCCGAGCTGACCTATTTGAG 1865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2020 ATGACATTTGAGATGATGCTTTTAAAGTACAGAGAGAAATGCAAGAAAGAGAG 2079
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1866 ATGAGATTTGAAATGATGCTGTGCTGACGTTTCAAGAGAAAGTGCAGAGAA 1925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2080 GGCCTTTGGGCTGTTAGAGACAGTGGACTATTCAGAACCAAAACCCAGAGAG 2139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1926 AATCTTTGAGCTTTGTGAGAGACAGGAGACCAACAGCAAGAAATAA---CT 1982
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2140 AAGAAATATCATTTATTTGAAATGTCTGCTCTCAACATCAAGAGCTTTAGAT 2199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1983 GAGGATTCGACACATTTTCAAGATTAACGATCAATTAATCTAGAGAACTTCA 2042
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2200 AAGAAACAGTTTGAAGAAATGTTGACAGTTCTAAAGGTGAGAGATGACAA 2259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2043 AAGAAACAGTTTGAAGAAATGTTGAGAGAAATGATGAGAGAGAGAGAGAG 2102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2260 GTTCTATGCTGCTCTTCCAAAGAGAAAGAAATATGAGAGAGAGAGAGAGAG 2319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2103 GTTCTATGCTGCTCTTCCAAAGAGAAAGAAATATGAGAGAGAGAGAGAGAG 2162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2320 CTGTGAGCATATGAGCTGTTTACAGAGTCCATTAACAGTTTGCAGATGATGAT 2379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2163 TCTGAGAGCCAAAGAGTTGTTTACAGAGTCCCATCATCATGTTTGCATATGAG 2222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2380 GTTGGCTTTCAAGAAATGTAATCTGACACCTTGGATTCAAATAATGAGAGCT 2439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```
Db 2223 GTCGCTTCCACAGAAATGATTCGACATCTATAACCCAGTTTATGAGCCGGCATATAT 2282
Qy 2440 TTCACCAAGAACTCAAAAACCTGGCCAGAGAAAGCCAAAGAAATCTGCTGCAGATAG 2499
Db 2283 TTCACCAAGAACTCAAAAACCTGACAGACAGCAAGAAACCTCAAGCACACAGACAG 2342
Qy 2500 CTGATCTGATGTTGAGGCTGAAATGATCTACAGAGCTCTCTGCGCAGGACATCCGTTA 2559
Db 2343 CTATCTCTGTTGTTGAGGCGAGAACTACTCACAGGCTCTCTGTCAGGATATCTCTCA 2402
Qy 2560 AATATTTGTCGCCACCACTGAGTCTGAGAGCTATAGATGTCATGACAGTGTGGTTGAC 2619
Db 2403 AATATCATCCCTCCACGATGAGTCTGAGGCTTAAATGTCATGACAGAGCTATGAGC 2462
Qy 2620 AATGCTCCAGCCCTGAAACCTTTTATTTTATGTCATGACAGGCTATACCTGATAT 2679
Db 2463 AATGTTCCAGCCCTGAAACCAATGTTGTTTATGTCATGACAGGCTATGCTGATAC 2522
Qy 2680 TTGTCAGATGACCCAGC-----AATATGTCAGTCACACAGATTAC 2721
Db 2523 TTGTCAGATGTCACACAGATGATGACATTCACAGCATCCGATGTGTCACAGACTAC 2582
Qy 2722 TCATCAGACCAATGAGACCCCTTTCACAGCATCCTTGAGGGATTCGCAAGTGCAGC 2781
Db 2583 TCATCAGACCAAGAAATGCTCTCTCGCTGCGATGCTGGAATGGCTCTTAATGAGCAGC 2642
Qy 2782 CCTGTGATTAATCTCTCATCATTTTAAACAGCTGTATGSCCTTACCTGGGTAACTA 2841
Db 2643 TCTGT---TTAGTGTCTCATCATGTTTAAACAAGAGG---GTTGAGAGACTG 2693
Qy 2842 ACCAATATATGACCATTCAGTGGCTCAAGAGAGTGCTGAATATCCCATGGTTATGTCG 2901
Db 2694 ACAAATGATATAAT-----ACAGCTTAACTG 2720
Qy 2902 TATGACTGACTGGGTTATGAAAGACTAGCCACATACATCAATTAAGTCCCTTATC 2961
Db 2721 TTCAGAAATGATGGGCTCACTAAAGCAGCCAGCACACACTAGCATCATAGTCCCTTTC 2780
Qy 2962 TGTCTTATGCTTGGGTTGGGGTAGATATACCAATGAACACTTTCAGACCTTC 3021
Db 2781 TTTACCTCT---GGGCTTGACAGGAGATGCCACTTAACCTTCTCAGCTGCTTT 2834
Qy 3022 CTTCCTCTGAGTGTCTTAAATCTCCTTACTAGAGAGATAAATATTTTGGATATA 3081
Db 2835 TCT-ATTGACATCTTTATCTCTTCTATAGGTGACAGCAAGAAATCTTATATAGA 2893
Qy 3082 ATGAAGAAATTTTCTAGTATATAAGCAGGCTTTTATTTTCTAAATGATGATAGTAT 3141
Db 2894 ACAAGGATATTTT---CAAGCTGTTATTTTCTAAATGATAGACAA 2942
Qy 3142 AAAATGTTAGATTAACAGAAATGATTTTTCAGAGAAATATATAAGTCTTTA 3201
Db 2943 C-----TAGACAACAGGATTTTCTATATAT-TTATTAAGTCTTTG 2994
Qy 3202 GGTATGAAATTAATCATCTTGTCTGAT 3231
Db 2995 GATATCCAATTAATCACCCTTGTCTGAGT 3024

RESULT 8
ABV23030
ID ABV23030 standard; cDNA: 854 BP.
XX
AC ABV23030;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 23021.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
```

```
OS Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
XX
XX 16-MAR-2000; 2000US-183862P.
XX
XX 25-MAY-2000; 2000US-207454P.
XX
XX 09-JUN-2000; 2000US-211314P.
XX
XX 18-JUL-2000; 2000US-219007P.
XX
XX 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI, 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer
XX
XX Claim 1; Page 4103-4104; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 854 BP; 272 A; 171 C; 183 G; 227 T; 1 other:
XX
XX Query Match 24.4%; Score 792; DB 23; Length 854;
XX Best Local Similarity 98.2%; Pred. No. 3.6e-205;
XX Matches 823; Conservative 0; Mismatches 10; Indels 5; Gaps 2;
Qy 898 GTCATCTATAAATACTACATTTAAGACAGTAGCAATTCAGCCTTGAGCTCTGGATT 957
Db 20 GTCATCTATAAATAATCTACATTTAAGACAGTAGCAATTCAGCCTTGAGCTCCGCG--- 76
Qy 958 TTTCAATCCCTCTGAATTTGTGTCAAAAGACTATTTAGACATATCCGGGTTAGTTTG 1017
Db 77 -CAGCGGTGGTCTGATTTGTGTACAAAGACTATTTGTAGACATATCCGGGTTAGTTTG 135
Qy 1018 CAAGGGAAGCAATGATGATTAATTTGAAAGAAATTCACCTGTGAGCAATAGACACCT 1077
Db 136 CAAGGGAAGCAATGATGATTTGAAAGAAATTCACCTGTGAGCAATAGACACCT 195
Qy 1078 ACTGTTGCTGCTTTAAAGCTCTTACAGAAATTCATCTTGGGAGAGTGGAGACAA 1137
Db 196 ACTGTTGCTGCTTTAAAGCTCTTACAGAAATTCATCTTGGGAGAGTGGAGACAA 255
Qy 1138 GAAACACCCCTTCTTCAATGCAATGCTGTGAACAACCTGACCCCTCAGATTGTCAG 1197
Db 256 GAAACACCCCTTCTTCAATGCAATGCTGTGAACAACCTGACCCCTCAGATTGTCAG 315
Qy 1198 GGCACATTTGAATGGCAGAGGAGAGATGTAATTTGTAATTTCTGTAACCCACATGATAT 1257
Db 316 GGCACATTTGAATGGCAGAGGAGAGATGTAATTTGTAATTTCTGTAACCCACATGATAT 375
```

QY 1258 ACAGTTGGACCTGTGGCAAGTCAATTTCTACAACAGCAGAGATTGAATGAAATCGGAA 1317
DB 376 ACAGTTGGACCTGTGGCAAGTCAATTTCTACAACAGCAGAGATTGAATGAAATCGGAA 435
QY 1318 TTTCTTGGCCAAAGGCTTAACAGTTTCAACGGTCCAGTTGGTACTGGTCACAAAAGGA 1377
DB 436 TTTCTTGGCCAAAGGCTTAACAGTTTCAACGGTCCAGTTGGTACTGGTCACAAAAGGA 495
QY 1378 TTTTACTTGTCTGTAATAATATATATACATGTACTGTGGCATTGAGAAATTTCTTAAACCT 1437
DB 496 TTTTACTTGTCTGTAATAATATATATACATGTACTGTGGCATTGAGAAATTTCTTAAACCT 555
QY 1438 CAAATATTAACATGATGAAAGAGTGTGGTGAATAATGCAATTGACAAATATTAAT 1497
DB 556 CAAATATTAACATGATGAAAGAGTGTGGTGAATAATGCAATTGACAAATATTAAT 615
QY 1498 TCCATTTCCTTCTGCTGCTGGAGCTGGAAACATGGAATAAAGAAAGCAAGAGCA 1557
DB 616 TCCATTTCCTTCTGCTGCTGGAGCTGGAAACATGGAATAAAGAAAGCAAGAGCA 675
QY 1558 GAGA-TTTTGTGTGATGAAGTTTAAACATTTGCCAAAGACATGTAAACACCAGTTAAC 1616
DB 676 GAGATTTTGTGTGATGAAGTTTAAACATTTGCCAAAGACATGTAAACACCAGTTAAC 735
QY 1617 TGTAAATTTTGTGATCTTTTCCAAACAGATTGGAGATATATAGGCTTTCAGTTTGAAT 1676
DB 736 TGTAAATTTTGTGATCTTTTCCAAACAGATTGGAGATATATAGGCTTTCAGTTTGAAT 795
QY 1677 GGCAAAGAGGTCCAAAGATGCTGAGTTTGAACATTTACAGTGTCCCGCAGTCAACCGCA 1734
DB 796 GGCAAAGAGGTCCAAAGATGCTGAGTTTGAACATTTACAGTGTCCCGCAGTCAACCGCA 853

RESULT 9

-ABV28866
ID ABV28866 standard; cDNA; 854 BP.

XX ABV28866;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker CDNA 28857.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05171.

XX 17-FEB-2000; 2000US-183319P.

XX 16-MAR-2000; 2000US-189862P.

XX 25-MAY-2000; 2000US-207454P.

XX 09-JUN-2000; 2000US-211314P.

XX 18-JUL-2000; 2000US-219007P.

XX 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

XX Claim 1; Page 6082-6083; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) determining the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker.

SO Sequence 854 BP; 272 A; 171 C; 183 G; 227 T; 1 other;

Query Match 24.4%; Score 792; DB 23; length 854;

Best Local Similarity 98.2%; Pred. No. 3.6e-205;

Matches 823; Conservative 0; Mismatches 10; Indels 5; Gaps 2;

QY 898 GTCATCTATTAATAATCTACATTAAGACAGTACGATTCAGCTTGGGATT 957

DB 20 GTCATCTATTAATAATCTACATTAAGACAGTACGATTCAGCTTGGGATT 76

QY 958 TTTCACTTCCCTCTGATTTTGTACAAAGACTATGTAGACATTCAGCTTGGGATT 1017

DB 77 -GACGGTGGGCTGTGATTTTGTACAAAGACTATGTAGACATTCAGCTTGGGATT 135

QY 1018 CAAAGGAGCCCAATGATGATTAATTTGAAAGAAATTCACCTGGTGAAGACACCT 1077

DB 136 CAAAGGAGCCCAATGATGATTAATTTGAAAGAAATTCACCTGGTGAAGACACCT 195

QY 1078 ACTGTGCTGCTTTTAACTGCTTCAAGATTCATCTGAGAGAGAGTACGCGGACCA 1137

DB 196 ACTGTGCTGCTTTTAACTGCTTCAAGATTCATCTGAGAGAGAGTACGCGGACCA 255

QY 1138 GAAACCAACCCCTTCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 1197

DB 256 GAAACCAACCCCTTCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 315

QY 1198 GGCACATTTGAATGCGAGCGGAGATGATTTTGTATTTGTATTAACCCAGATGAT 1257

DB 316 GGCACATTTGAATGCGAGCGGAGATGATTTTGTATTTGTATTAACCCAGATGAT 375

QY 1258 ACAGTTGGACCTGTGGCAAGTCAATTTCTACAACAGCAGAGATTGAATGAAATCGGAA 1317

DB 376 ACAGTTGGACCTGTGGCAAGTCAATTTCTACAACAGCAGAGATTGAATGAAATCGGAA 435

QY 1318 TTTCTTGGCCAAAGGCTTAACAGTTTCAACGGTCCAGTTGGTACTGGTCACAAAAGGA 1377

DB 436 TTTCTTGGCCAAAGGCTTAACAGTTTCAACGGTCCAGTTGGTACTGGTCACAAAAGGA 495

QY 1378 TTTTACTTGTCTGTAATAATATATATACATGTACTGTGGCATTGAGAAATTTCTTAAACCT 1437

DB 496 TTTTACTTGTCTGTAATAATATATATACATGTACTGTGGCATTGAGAAATTTCTTAAACCT 555

QY 1438 CAAATATTAACATGATGAAAGAGTGTGGTGAATAATGCAATTGACAAATATTAAT 1497

DB 556 CAAATATTAACATGATGAAAGAGTGTGGTGAATAATGCAATTGACAAATATTAAT 615

QY 1498 TCCATTTCCTTCTGCTGCTGGAGCTGGAAACATGGAATAAAGAAAGCAAGAGCA 1557

DB 616 TCCATTTCCTTCTGCTGCTGGAGCTGGAAACATGGAATAAAGAAAGCAAGAGCA 675

QY 1558 GAGA-TTTTGTGTGATGAAGTTTAAACATTTGCCAAAGACATGTAAACACCAGTTAAC 1616

DB 676 GAGATTTTGTGTGATGAAGTTTAAACATTTGCCAAAGACATGTAAACACCAGTTAAC 735

QY 1617 TGTAAATTTTGTGATCTTTTCCAAACAGATTGGAGATATATAGGCTTTCAGTTTGAAT 1676

DB 736 TGTAAATTTTGTGATCTTTTCCAAACAGATTGGAGATATATAGGCTTTCAGTTTGAAT 795

```
QY 1677 GCCTAAGAGGTCCAGATGCTGATTGCAACAATTACAGTGTCCCGCAGTCAACGAGA 1734
    |||||||
Db 796 GGCCTAAGAGGTCCAGATGCTGATTGCAACAATTACAGTGTCCCGCAGTCAACGAGA 853

RESULT 10
AAH72506
ID AAH72506 standard; cDNA; 822 BP.
XX
AC AAH72506;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 3780.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN W0200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US33312.
XX
PR 08-DEC-1999; 99US-0169681.
PR 21-DEC-1999; 99US-0171350.
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer
PT and for assessing and detecting compounds for treating the cancer -
XX
PS Claim 1; Page 718; 1051pp; English.
XX
CC The invention relates to novel genes (AAH6872-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy.
XX
SQ Sequence 822 BP; 228 A; 177 C; 211 G; 198 T; 8 other;

Query Match 19.2%; Score 622.8; DB 22; Length 822;
Best Local Similarity 86.18; Pred. No. 4.5e-159;
Matches 759; Conservative 0; Mismatches 17; Indels 106; Gaps 2;

QY 121 GCGGGAGCGGGCGGAGAGTGGGCGCCATATCTGGAACACTATGCTTATGCTTGA 180
    |||||
Db 45 GAGGAGACCGGCGCTGCCGGAAGTGGGCGCCATATCTGGAACACTATGCTTATGCTTGA 104

QY 181 AGCCGCAAAAGGGAATTAACATTTAAAGACTCCCGGGGACCTGGAGATGGAATTTTCC 240
    |||||||
Db 105 AGCCGCAAAAGGGAATTAACATTTAAAGACTCCCGGGGACCTGGAGATGGAATTTTCC 164

QY 241 ATGTGTGCGCGGAGAGCAAGCTTACATTAATAAATCAGGTAGGATTTACTCGCTCTCACTC 300
    |||||||
Db 165 ATGTGTGCGCGGAGAGCAAGCTTACATTAATAAATCAGGTAGGATTTACTCGCTCTCACTC 197

QY 301 TTGTTTCAAGAAAGTCTTGTCTCAGATCTTCTCAGTGAAGGAGGGAATACAGAAAGA 360
```

```
Db 198 ----- 197
QY 361 TGTCTCCCTACAGATGCTCAGAGAGCTGCTGCTGTGGAGAAACATPATGTGGCAAT 420
    |||||||
Db 198 -----TACGAGACTGTGCTCTTGTGAGAAACATPATGTGGCAAT 239

QY 421 CCCATTAAACCAATTAAGCTTCAAAATTTTAAATAATATGAGCTCAGCTGTGTGAATC 480
    |||||||
Db 240 CCCATTAAACCAATTAAGCTTCAAAATTTTAAATAATATGAGCTCAGCTGTGTGAATC 299

QY 481 CTCGAGATTAAGTTTGGCTGTATCTTAACCTGATCTCCAGTTCAGGAAGCAAGC 540
    |||||||
Db 300 CTCGAGATTAAGTTTGGCTGTATCTTAACCTGATCTCCAGTTCAGGAAGCAAGC 359

QY 541 AAATCTCTGCAAGTCTCAGAAATAATGCTGCTCTGTGAGATGATATCATGCTGAAA 600
    |||||||
Db 360 AAATCTCTGCAAGTCTCAGAAATAATGCTGCTCTGTGAGATGATATCATGCTGAAA 419

QY 601 GATGACCTCACACACATGCTGTGATGCTGTGATGCTGATGCAAGCAATGAAGATCTCTG 660
    |||||||
Db 420 GATGACCTCACACACATGCTGTGATGCTGTGATGCTGATGCAAGCAATGAAGATCTCTG 479

QY 661 CATGGGGAGGCTGTGGCCCTGGCCCTGTAAAGCTGTGATTTGAATTCAGAAAG 720
    |||||||
Db 480 CATGGGGAGGCTGTGGCCCTGGCCCTGTAAAGCTGTGATTTGAATTCAGAAAG 539

QY 721 AGCAACAAGTTTGTGGCAGATATGTAAGTGTACACTGCTGTGATATGCTGTACGGGA 780
    |||||||
Db 540 AGCAACAAGTTTGTGGCAGATATGTAAGTGTACACTGCTGTGATATGCTGTACGGGA 599

QY 781 GCAGGAGAGGCTTCCCTGCAACACATCATCTCTGTGGGCTCGGTGATGGAATG 840
    |||||||
Db 600 GCAGGAGAGGCTTCCCTGCAACACATCATCTCTGTGGGCTCGGTGATGGAATG 659

QY 841 GATAAACAGGAGTACTGCAAAAGCTCAGAGGG-CCATTGTAGTATTTCTGAATATGT 899
    |||||||
Db 660 GATAAACAGGAGTACTGCAAAAGCTCAGAGGGCCCATTTGTATGTCGGAATATGT 719

QY 900 CATCTATTAATAATCTCATATTAGACAGTAGCAATTCACACCTTGAGCTTGGGATTT 959
    |||||||
Db 720 CATCTATTAATAATCTCATATTAGACAGTAGCAATTCACACCTTGAGCTTGGGATTT 779

QY 960 TCAGTCCCTCTGAATTTGTGTACAAAGACTATTTGTAGAGAC 1001
    |||||||
Db 780 TCAGTCCCTCTGAATTTGTGTACAAAGACTATTTGTAGAGAC 821

RESULT 11
ABV30165
ID ABV30165 standard; cDNA; 820 BP.
XX
AC ABV30165;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 30156.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
```

```
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX WPI: 2001-662795/76.
DR
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
PS
XX Claim 1; Page 6531; 11750pp; English.
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 820 BP; 226 A; 178 C; 211 G; 197 T; 8 other;
Query Match 18.8%; Score 609; DB 23; Length 820;
Best Local Similarity 85.1%; Pred. No. 2,6e-155;
Matches 750; Conservative 0; Mismatches 25; Indels 106; Gaps 2;
QY 121 GCGGGAGCGGCGGCGCAGAGTGGGCCACATATCTGGAAACCTACAGTCTATGCTTTGA 180
DB 45 GAGGAGAGCGGCGCTCGCGGAGTGGGCCACATATCTGGAAACCTACAGTCTATGCTTTGA 104
QY 181 ACCGCAAAAGGGAATTAACATTTAAAGACTCCCGGGGAGCTGGAGAGATGAGACTTTTCC 240
DB 105 ACCGCAAAAGGGAATTAACATTTAAAGACTCCCGGGGAGCTGGAGAGATGAGACTTTTCC 164
QY 241 ATGTGGCCCGGAGCAGACACTTACAATGAAAAATCAGTAGATTACTCGCTCCTCAGTC 300
DB 165 ATGTGGCCCGGAGCAGACACTTACAATGAAAAATCAGTAGATTACTCGCTCCTCAGTC 197
QY 301 TTGTTTCAGAAAGTCTTTGCTCAGATCTTTCTCTCAGTGGAGAAAGGGAATACAGAAAGA 360
DB 198 ----- 197
QY 361 TGTCTCCCTACAGAGTGTCTCAGAGACTGTGCTTTGGAAAACTATATGTTGGCAATT 420
DB 198 -----TCAGAGACTGTGCTTTGGAAAACTATATGTTGGCAATT 239
QY 421 CCCATTAAACCAATGACTCTTCAAAATTTTAAAAAATTAATGAGCGAGCTGTGTAAGTC 480
DB 240 CCCATTAAACCAATGACTCTTCAAAATTTTAAAAAATTAATGAGCGAGCTGTGTAAGTC 299
QY 481 CTCAGAAATAGTTGGCTGTATCTCTACCCCTGGTCTCCAGTTCAGAGAGCAACAGC 540
DB 300 CTCAGAAATAGTTGGCTGTATCTCTACCCCTGGTCTCCAGTTCAGAGAGCAACAGC 359
QY 541 AAATCTCTCAAGTGTTCGAAAAAATGCTGACTCCTAGATAGAGTTATCATGCTTGGAAA 600
DB 360 AAATCTCTCAAGTGTTCGAAAAAATGCTGACTCCTAGATAGAGTTATCATGCTTGGAAA 419
QY 601 GATGAGCTACACACATGCTGTGATGCTGTGGGGAATGAGCAAAATGAAGTCTTTCG 660
DB 420 GATGAGCTACACACATGCTGTGATGCTGTGGGGAATGAGCAAAATGAAGTCTTTCG 479
QY 661 CATGGGGAGGCGCTGGCCCTGGCTGATAAAGCTGTGATTGTAATCCAGAGAGAG 720
```

```
DB 480 CATGGGGAGGCGCTGGCCCTGGCTGATAAAGCTGTGATTGTAATCCAGAGAG 539
QY 721 ACCAAACGTTTGTTCGAGATATGTAAGTGTGAGTGTGAGATAGCTTCACGGGA 780
DB 540 ACCAAACGTTTGTTCGAGATATGTAAGTGTGAGTGTGAGATAGCTTCACGGGA 599
QY 781 GCAGGAGGCTTCCCTGCAACAGATATCATGCTGTGGGCGCTCGGTGATGGAATGG 840
DB 600 GCAGGAGGCTTCCCTGCAACAGATATCATGCTGTGGGCGCTCGGTGATGGAATGG 659
QY 841 GATTAACAGGATGTACTGGAAGAGCTGCAGAGAGGCCCATTTGATATTCGATTATGTC 900
DB 660 GATTAACAGGATGTACTGGAAGAGCTGCAGAGAGGCCCATTTGATATTCGATTATGTC 718
QY 901 ATCTATTAATAATACATACATTAAGACAGTAGCAATTCAGCCTTGAGCTGTGGAATTTT 960
DB 719 ATCTATTAATAATACATACATTAAGACAGTAGCAATTCAGCCTTGAGCTGTGGAATTTT 778
QY 961 CAGTTCCTCTGCAATTTGTGTACAAAGACTATTTAGAGAGC 1001
DB 779 CAGTTCCTCTGCAATTTGTGTACAAAGACTATTTAGAGAGC 819
```

```
RESULT 12
AAH72659
ID AAH72659 standard; cDNA; 666 BP.
XX
AC AAH72659;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 3933.
XX
KM Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US33312.
XX
PR 08-DEC-1999; 99US-0169681.
PR 21-DEC-1999; 99US-0171350.
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI: 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer
PT and for assessing and detecting compounds for treating the cancer -
PS
XX Claim 1; Page 785; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy.
XX
SQ Sequence 666 BP; 191 A; 139 C; 183 G; 152 T; 1 other;
```

Query Match	15.1%	Score 489.4	DB 22	Length 666
Best Local Similarity	84.6%	Pred. No. 9.1e-123		
Matches 608	Conservative 0	Mismatches 6	Indels 105	Gaps 1
QY 136	GCACAAAGTGGGCCACCAATATCTGTGAAGAACTACAGTATATGCTTTGAAGCGCAAAAGGAAAT	195		
Db 52	GA AAAAGTGGGCCACCAATATCTGTGAAGAACTACAGTATATGCTTTGAAGCGCAAAAGGAAAT	111		
QY 196	AAACATTTTAAAGACTCCCCCGGGGACCTGGAGGATGACCTTTTCATGGTGGCCGGAGCA	255		
Db 112	AAACATTTTAAAGACTCCCCCGGGGACCTGGAGGATGACCTTTTCATGGTGGCCGGAGCA	171		
QY 256	GCAGCTTACATGA AAAATCAAGTAGGATTAACCTGCTCTCACTCTTGTTCAGAAAGTC	315		
Db 172	GCAGCTTACATGA AAAA-----	189		
QY 316	TTTGCTCAGATCTTCCCTCAGTGGAGAAAGGGGAATACAGAGAAATGTCCCTTACAG	375		
Db 190	-----	189		
QY 376	TGCTCAGAGACGTGCTCTTGGAGAAAACATTAAGTTGGCAATTCCTTAACCAAT	435		
Db 190	---TCAGGACTGGGCTCTTGGAGAAAACATTAAGTTGGCAATTCCTTAACCAAT	246		
QY 436	GACTTCAAAATTTTAAAAATATAGGCGCTGCTGTGAAGTCCCTCAGAAATAAGTT	495		
Db 247	GACTTCAAAATTTTAAAAATATATAGGCGCTGCTGTGAAGTCCCTCAGAAATAAGTT	306		
QY 496	GGCTGTATCTTACCCCTGCTCTCCAGTTCAAGGAAAGCAACAGCAATCTTGCAGTG	555		
Db 307	GGCTGTATCTTACCCCTGCTCTCCAGTTCAAGGAAAGCAACAGCAATCTTGCAGTG	366		
QY 556	TTACAGAAAATGCTACTCCGTAGATAGATAGTATCAGTCTGGAAAAGATGACCTACACACA	615		
Db 367	TTACAGAAAATGCTACTCCGTAGATAGATAGTATCAGTCTGGAAAAGATGACCTACACACA	426		
QY 616	CATGCTGTGTGATGCTGTGGTGAATCAGCCCAATGAAGATCTTCTCATGGGGAGGCTCG	675		
Db 427	CATGCTGTGTGATGCTGTGGTGAATCAGCCCAATGAAGATCTTCTCATGGGGAGGCTCG	486		
QY 676	GCCCTGGCCCTGGTAAAAAGCTGTGGATTGAAATCCAGAAGAGAGCAACAGTTGTT	735		
Db 487	GCCCTGGCCCTGGTAAAAAGCTGTGGATTGAAATCCAGAAGAGAGCAACAGTTGTT	546		
QY 736	GCCAAATATGGAAGTGTACGCTGGTGAATGAGCTGCAGGGGAGCAGGAGGCTTCCC	795		
Db 547	GCCAAATATGGAAGTGTACGCTGGTGAATGAGCTGCAGGGGAGCAGGAGGCTTCCC	606		
QY 796	TGCAAAACAGATCATCATGCTGTTGGGCTCGGTGGATGGAATGGGATAAACAGGATG	854		
Db 607	TGCAAAACAGATCATCATGCTGTTGGGCTCGGTGGATGGAATGGGATAAACAGGATG	665		
RESULT 13				
AAH70240/c				
AAH70240 standard; cDNA; 516 BP.				
AAH70240;				
19-SEP-2001 (first entry)				
Human cervical cancer marker nucleic acid 1514.				
Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.				
Homo sapiens.				
MO200142467-A2.				
14-JUN-2001.				
08-DEC-2000; 2000MO-US33312.				

PR 08-DEC-1999; 99US-0169681.
PR 21-DEC-1999; 99US-0171350.
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
XX WPI: 2001-375006/39.
XX
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer
PT and for assessing and detecting compounds for treating the cancer -
XX
XX
PS Claim 1; Page 341; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy.
XX
XX Sequence 516 BP; 124 A; 138 C; 108 G; 146 T; 0 other;

[illegible]

```
XX Human cervical cancer marker nucleic acid 2944.
DE Cervical cancer: cytostatic; pre-malignant condition; gene therapy. ss.
XX
XX Homo sapiens.
XX MO200142467-A2.
XX
XX 14-JUN-2001.
XX
XX 08-DEC-2000; 2000WO-US33312.
XX
XX 08-DEC-1999; 99US-0169681.
XX 21-DEC-1999; 99US-0171350.
XX 14-MAR-2000; 2000US-0189315.
XX 12-MAY-2000; 2000US-0203791.
XX 09-JUN-2000; 2000US-0210600.
XX 21-JUL-2000; 2000US-0220114.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer
XX and for assessing and detecting compounds for treating the cancer -
XX
XX Claim 1; Page 597; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
XX polypeptides are useful: to assess if a patient is afflicted with
XX cervical cancer or has a pre-malignant condition; to monitor the
XX progression of cervical cancer or a premalignant condition in a patient;
XX and to select and/or assess the efficacy of a compound or therapy for
XX inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy.
XX
XX Sequence 481 BP; 116 A; 129 C; 102 G; 134 T; 0 other;
XX
Query Match 13.9%; Score 449.8; DB 22; Length 481;
Best Local Similarity 99.4%; Pred. No. 4.7e-112;
Matches 462; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 394 CTTGGAGAACTAGTGTGGCAAA-TTCCCATTAACCAATGACTTCAAAATTTTAA 452
DB 481 CTTGGAGAACTAGTGTGGCAAAATTTCCCATTAACCAATGACTTCAAAATTTTAA 422
QY 453 AAATATGAGCGTCAGCTGTGAAGTCTCCAGATAAGTTGGCTATCTACCCCT 512
DB 421 CAATATATGAGCGTCAGCTGTGAAGTCTCCAGATAAGTTGGCTATCTACCCCT 362
QY 513 GGTCTCTCCAGTTCAGGAAGCAACAGCAAAATCTCTGCAAGTGTTCAGAAAATGCTGAC 572
DB 361 GGTCTCTCCAGTTCAGGAAGCAACAGCAAAATCTCTGCAAGTGTTCAGAAAATGCTGAC 302
QY 573 TCCATGAGTATAGTATAGTCTGGAAGATGACCTCACCACATGCTGTGATGCTGT 632
DB 301 TCCATGAGTATAGTATAGTCTGGAAGATGACCTCACCACATGCTGTGATGCTGT 242
QY 633 GGTGAATCAGCCCAATGAAGATCTTGTGATGGGAGAGCCCTGGCCCTGGGTAA 692
DB 241 GGTGAATCAGCCCAATGAAGATCTTGTGATGGGAGAGCCCTGGCCCTGGGTAA 182
QY 693 AGCTGGTGGATTTGAATCAAGAAGAGAGCAAAAGTTTGTGCGATATGTAAGT 752
DB 181 AGCTGGTGGATTTGAATCAAGAAGAGAGCAAAAGTTTGTGCGATATGTAAGT 122
QY 753 GTCAGCTGGTATAGTGTTCACGGGAGCAGGAGGCTCCCTGCAAAACGATCATCA 812
DB 121 GTCAGCTGGTATAGTGTTCACGGGAGCAGGAGGCTCCCTGCAAAACGATCATCA 62
```

```
QY 813 TGTCTTGGCCCTCGGTGATGATGGGATGAATGAACAGGATGTAC 857
DB 61 TGTCTTGGCCCTCGGTGATGATGGGATGAATGAACAGGATGTAC 17
RESULT 15
AB056473/c
ID AB056473 standard; cDNA; 668 BP.
XX
XX AC AB056473;
XX
XX 02-AUG-2002 (first entry)
XX
XX Human colon cancer related nucleotide sequence SEQ ID NO:168.
XX
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX MO200229086-A2.
XX
XX 11-APR-2002.
XX
XX 02-OCT-2001; 2001WO-US30732.
XX
XX 02-OCT-2000; 2000US-237271P.
XX
XX (FARB ) BAYER CORP.
XX
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
XX Thlaglingam A, Lewis ME;
XX WPI; 2002-426115/45.
XX
XX New isolated nucleic acid that is differentially expressed in cancer
XX tissues useful for determining the presence of colon cancer in a cell
XX or tissue type, and in antisense therapy -
XX
XX Claim 1; Fig 1; 796pp; English.
XX
XX AB056306 to AB060787 represent isolated nucleic acids (I) differentially
XX expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
XX encoded by the AB060776 to AB060787 nucleic acid sequences. (I) can be
XX used in antisense therapy. An antibody immunoreactive with a polypeptide
XX encoded by (I) is useful for detecting cancer in a patient sample, and
XX for detecting the presence or absence of (I) in a cell. A probe/primer derived
XX from (I) can be used for determining the presence of a nucleic acid which
XX hybridises to (I), and for determining the phenotype of cells in a sample
XX of cells from a patient. (I) is useful for determining the presence of
XX colon cancer in a cell or tissue type, for determining the presence of
XX state of other type of cancer, in antisense therapy, to generate
XX macroarrays on a solid surface, to identify a chromosome on which the
XX corresponding gene resides, and in tissue profiling, forensics, genetic
XX analysis, mapping and diagnostic applications. (I) can be used to raise
XX antibodies, and to screen for peptide analogues and antagonists.
XX
XX Sequence 668 BP; 145 A; 152 C; 137 G; 175 T; 59 other;
XX
Query Match 13.1%; Score 424.2; DB 24; Length 668;
Best Local Similarity 96.5%; Pred. No. 5.3e-105;
Matches 462; Conservative 0; Mismatches 14; Indels 3; Gaps 3;
QY 379 TCAGAGACTGTGCTCTTGGAGAAACTATAGTGGCAATTTCCATTACCAATGAC 438
DB 477 TCAAGATGAGTGTCTTGAAGAAACCTATAGTGGCAATTTCCATTACCAATGAC 418
QY 439 TTCAAAATTTTAAATAATGAGCGTCAGCTGTGTGAAGTCTCCAGATAAGTTGGC 498
DB 417 TTCAAAATTTTAAATAATGAG-GTCAGCTGTGTGAAGTCTCCAGATAAGTTGGC 359
```

```
QY 499 TGTATCTCTACCCCTGGTCTCTCCAGTTCAGGAAGCACAACAATCTCTGCAAGTGTTC 558
  |||||||
Db 358 TG-ATCTCTACCCCTGGTCTCTCCAGTTCAGGAAGCANCACAATCTCTGCAAGTGTTC 300
  |||||||
QY 559 AGAAAAATGCTGACTCCTAGGATAGAGTATTCAGTCTGGAAGATGACTCACCACACAT 618
  |||||||
Db 299 AG-AAAAATGCTGACTCCTAGGATAGAGTATTCAGTCTGGAAGATGACTCACCACACAT 241
  |||||||
QY 619 GCTGTGATGCTGTGGATGCAATGAGCATCTTGCATGGGGAGGCTTGCC 678
  |||||||
Db 240 GCTGTGATGCTGTGGATGCAATGAGCATCTTGCATGGGGAGGCTTGCC 181
  |||||||
QY 679 CTGGCCCTGCTAAAGCTGCTGATTTGAATCCAAAGAGAGCAACAGTTTGTGCC 738
  |||||||
Db 180 CTGGCCCTGCTAAAGCTGCTGATTTGAATCCAAAGAGAGCAACAGTTTGTGCC 121
  |||||||
QY 739 AGATATGCTAAAGCTGCTGAGATAGCTGTCACGGGAGCAGGAGCTTCCCTGC 798
  |||||||
Db 120 AGATATGCTAAAGCTGCTGAGATAGCTGTCACGGGAGCAGGAGCTTCCCTGC 61
  |||||||
QY 799 AAACAGATCATCATGCTGTGGGCTCGGTGATGGAATGGAATAACAGGATGTAC 857
  |||||||
Db 60 AAACAGATCATCATGCTGTGGGCTCGGTGATGGAATGGAATAACAGGATGTAC 2
```

Search completed: February 7, 2003, 02:08:44
Job time : 452 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 04:00:54 : Search time 112 Seconds
(Without alignments)
13724.926 Million cell updates/sec

Title: US-09-830-762-1
Perfect score: 3243
Sequence: 1 gggtctcgcttcctcggtg.....gtctgattaaaaaaaaa 3243

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2842.6	87.7	3016	10 US-09-882-529-1	Sequence 1, Appl
2	299	9.2	299	10 US-09-920-300A-946	Sequence 946, App
3	299	9.2	299	12 US-10-033-528-946	Sequence 946, App
4	284	8.8	284	10 US-09-962-436-149	Sequence 149, App
5	246.8	7.6	260	10 US-09-964-824A-308	Sequence 308, App
6	226	7.0	238	9 US-09-796-692-7357	Sequence 7357, App
7	202.8	6.3	214	9 US-10-076-622-24	Sequence 24, Appl
8	202.8	6.3	214	10 US-09-604-287A-24	Sequence 24, Appl
9	202.8	6.3	214	10 US-09-339-338-24	Sequence 24, Appl
10	202.8	6.3	214	12 US-10-007-805-24	Sequence 24, Appl
11	180	5.6	180	9 US-10-046-935-1590	Sequence 1590, App
12	180	5.6	180	9 US-09-878-178-1590	Sequence 1590, App
13	140.4	4.3	2762	10 US-09-957-635-1	Sequence 1, Appl
14	140	4.3	2217	10 US-09-957-635-3	Sequence 1, Appl
15	121.2	3.7	3348	12 US-10-044-090-664	Sequence 664, App
16	70.2	2.2	71	10 US-09-998-598-2276	Sequence 2276, App
17	44.8	1.4	533	10 US-10-015-219-138	Sequence 138, App
18	44.8	1.4	533	10 US-09-777-564-138	Sequence 138, App
19	44.8	1.4	649	9 US-10-015-219-375	Sequence 375, App

20	44.8	1.4	649	10 US-09-777-564-375	Sequence 375, App
21	42.4	1.3	1992	10 US-09-864-761-10189	Sequence 10189, A
22	41	1.3	1136	9 US-09-938-842A-3287	Sequence 3287, App
23	40	1.2	127197	9 US-09-754-853A-1	Sequence 1, Appl
24	39.6	1.2	408	10 US-09-960-352-6263	Sequence 6263, App
25	39.6	1.2	2000	9 US-09-938-842A-4600	Sequence 4600, App
26	39.4	1.2	229	10 US-09-969-373-1347	Sequence 1347, App
27	39.4	1.2	302250	10 US-09-962-832-154	Sequence 154, App
28	39.4	1.2	640681	10 US-09-790-988-1	Sequence 1, Appl
29	39.2	1.2	2672	12 US-10-044-090-699	Sequence 699, App
30	39	1.2	650	10 US-09-879-536-189	Sequence 189, App
31	39	1.2	513509	9 US-09-754-853A-4	Sequence 4, Appl
32	38.4	1.2	601	10 US-09-820-002-13	Sequence 13, Appl
33	38.4	1.2	2073	10 US-09-815-242-8362	Sequence 8362, App
34	38.4	1.2	2109	10 US-09-815-242-4555	Sequence 4555, App
35	38.4	1.2	21784	10 US-09-820-002-3	Sequence 3, Appl
36	38.2	1.2	640681	10 US-09-790-988-1	Sequence 1, Appl
37	38	1.2	237	10 US-09-880-107-11727	Sequence 11727, App
38	38	1.2	435	10 US-09-864-761-20804	Sequence 20804, App
39	38	1.2	447	10 US-09-864-761-4048	Sequence 4048, App
40	38	1.2	1089	10 US-09-925-302-301	Sequence 301, App
41	38	1.2	2184	9 US-09-880-192-31	Sequence 31, Appl
42	37.8	1.2	401	9 US-09-946-607-678	Sequence 678, App
43	37.8	1.2	401	10 US-09-795-668-678	Sequence 678, App
44	37.8	1.2	401	10 US-09-795-668-678	Sequence 678, App
45	37.8	1.2	406	10 US-09-960-352-10265	Sequence 10265, A

ALIGNMENTS

RESULT 1
US-09-882-529-1
Sequence 1, Application US/09882529
Patent No. US20020132317A1
GENERAL INFORMATION:
APPLICANT: Peyman, John A
APPLICANT: da Silva, Antonio
TITLE OF INVENTION: NOVEL INTERFERON-INDUCED TETRASPAN PROTEIN AND NUCLEIC
FILE REFERENCE: 15966-771
CURRENT APPLICATION NUMBER: US/09/882, 529
CURRENT FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: 60/211, 565
PRIOR FILING DATE: 2000-06-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 3016
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (123)..(2579)
US-09-882-529-1

Query Match 87.7%: Score 2842.6; DB 10; Length 3016;
Best local similarity 96.3%: Pred. No. 0;
Matches 2963; Conservative 0; Mismatches 9; Indels 105; Gaps 1;

QY 151 CATATCTGGAACATCACTATCTTGAAGCCCAAAAGGAAATTAACATTAAAGCT 210
DB 45 CAGGCTGGAACATCACTATCTTGAAGCCCAAAAGGAAATTAACATTAAAGCT 104
QY 211 CCCCCGGGACCTGAGAGATGAGCTTTTCATGTGGCCGAGCAGCAGCTTAAATGAA 270
DB 105 CCCCCGGGACCTGAGAGATGAGCTTTTCATGTGGCCGAGCAGCAGCTTAAATGAA 164
QY 271 AAATCAGTAGGATTAACCTGCTCTCCTCCTCTTGTTCAGAAAGTCTTGTCAAGATCTTT 330
DB 165 AAA----- 167

OY 331 CCTCAGTGAGAAAAGGGAATACAGAAAGTGTCTCCCTACAGTGTCTCAGAGACTGCT 390
Db 168 -----TCAGAGACTGCT 179
OY 391 GCTCTTGAGAAAACATATAGTTGGCAAAATCCCATTTAACCAAAATGACTTCAAAATTTTA 450
Db 180 GCTCTTGAGAAAACATATAGTTGGCAAAATCCCATTTAACCAAAATGACTTCAAAATTTTA 239
OY 451 AAAAATAATGAGCGTCAGCTGTGTGAAGTCCTCCAGAAATAGTTGGCTGTCTCTTACC 510
Db 240 AAAAATAATGAGCGTCAGCTGTGTGAAGTCCTCCAGAAATAGTTGGCTGTCTCTTACC 299
OY 511 CTGTCTCTCCAGTTTCAGGAAGCAACAGCAAAATCTCTCAGTAGTTTCAGAAAAATGCTG 570
Db 300 CTGTCTCTCCAGTTTCAGGAAGCAACAGCAAAATCTCTCAGTAGTTTCAGAAAAATGCTG 359
OY 571 ACTCCTAGAGTAGATTATCAGTCTTGGAAGAATGACCTCACACACATGCTGTGATGCT 630
Db 360 ACTCCTAGAGTAGATTATCAGTCTTGGAAGAATGACCTCACACACATGCTGTGATGCT 419
OY 631 GTGTGAATGACGACCAATGAGATCTTTCATGGGGGAGGCTGGCCCTGGCCCTGGTA 690
Db 420 GTGTGAATGACGACCAATGAGATCTTTCATGGGGGAGGCTGGCCCTGGCCCTGGTA 479
OY 691 AAAGCTGTGATTTGAAAATCCAGAGAGAGCAAAACAGTTTGTGCCAGATATGATAA 750
Db 480 AAAGCTGTGATTTGAAAATCCAGAGAGAGCAAAACAGTTTGTGCCAGATATGATAA 539
OY 751 GTGTGAGCTGTGAGATAGCTGTACAGGAGCAGGAGGCTTCCCTGCAGATATGCTATC 810
Db 540 GTGTGAGCTGTGAGATAGCTGTACAGGAGCAGGAGGCTTCCCTGCAGATATGCTATC 599
OY 811 CATGCGTTGGGCGCTGGTGGATGGAATGGGATAAACAGGAGTGTCTGGAAGGCTGCAG 870
Db 600 CATGCGTTGGGCGCTGGTGGATGGAATGGGATAAACAGGAGTGTCTGGAAGGCTGCAG 659
OY 871 AGGGCATTGTAGATATTCATATGATCTATATAAAATACACATTAAGACAGTA 930
Db 660 AGGGCATTGTAGATATTCATATGATCTATATAAAATACACATTAAGACAGTA 719
OY 931 GCAATTCACGCTTGAGCTCTGGGATTTTTCAGTTCCCTGTGAATTTGTGTACAAAGACT 990
Db 720 GCAATTCACGCTTGAGCTCTGGGATTTTTCAGTTCCCTGTGAATTTGTGTACAAAGACT 779
OY 991 ATTGTAGAGACTATCCGGGTAGTTTGCAGGAGCAACCAATGATGAGTATTTGAAAGAA 1050
Db 780 ATTGTAGAGACTATCCGGGTAGTTTGCAGGAGCAACCAATGATGAGTATTTGAAAGAA 839
OY 1051 ATTCACTGTGTAGAGCAATGAGAACCCCTACTGTTGCTGCCCTTTAAAGCTGCTCAGAAATTC 1110
Db 840 ATTCACTGTGTAGAGCAATGAGAACCCCTACTGTTGCTGCCCTTTAAAGCTGCTCAGAAATTC 899
OY 1111 ATTCCTAGGAAAGTAGAGCTGGGACAAAGAACCCCTCTTCTTTCATATGCAATGCTGCTG 1170
Db 900 ATTCCTAGGAAAGTAGAGCTGGGACAAAGAACCCCTCTTCTTTCATATGCAATGCTGCTG 959
OY 1171 AACCACTGACCTCCAGATTGTCAGGGCCACATGAAATGGCAGAGCAGATGTAAT 1230
Db 960 AACCACTGACCTCCAGATTGTCAGGGCCACATGAAATGGCAGAGCAGATGTAAT 1019
OY 1231 GTTAAATCTGTAAACCACATGATATTTACAGTTGACCTGTGGCAAAAGTCAATTTACAA 1290
Db 1080 CAAGCGAGGAGTTGAATGAAATGCGAATTTCTTGCCACAAAGGCTTAAAGTTTCAACCG 1139
OY 1291 CAAGCGAGGAGTTGAATGAAATGCGAATTTCTTGCCACAAAGGCTTAAAGTTTCAACCG 1350
Db 1351 TCCCAATTTGCTACTGCTCAAAAGAGTTTAACTTGTCTGTATAATATATATACCTGTA 1410
OY 1410 TCCCAATTTGCTACTGCTCAAAAGAGTTTAACTTGTCTGTATAATATATATACCTGTA 1199
OY 1411 CTGTGCAATTCAGAAATTTCTTAAACCTCAGATATTTAAACATGCAATGAAGAGTGTG 1470

Db 1200 CTGTGCAATTCAGAAATTTCTTAAACCTCAGATATTTAAACATGCAATGAGAGTGTG 1259
OY 1471 GAAAAATGATTTGACCAAAATATTAATCTCCATTTCCCTTCCGCCCCGCGCTGGAGTGAAC 1530
Db 1260 GAAAAATGATTTGACCAAAATATTAATCTCCATTTCCCTTCCGCCCCGCGCTGGAGTGAAC 1319
OY 1531 ATGGAATTAAGAGAAAGAAACACAGCAGAGATTTGTTTGATGAAGTTTAAATTTGCC 1590
Db 1320 ATGGAATTAAGAGAAAGAAACACAGCAGAGATTTGTTTGATGAAGTTTAAATTTGCC 1379
OY 1591 AAAGACATGTAAACACACAGTTAACTGTAAATTTGTGATCTTTCACACAGATTTGGAG 1650
Db 1380 AAAGACATGTAAACACACAGTTAACTGTAAATTTGTGATCTTTCACACATTTGGAG 1439
OY 1651 ATATTAAGGCTTTGATTTCAAAATGGCAAGAGCTCCACATGCTGATTTGAACAT 1710
Db 1440 ATATTAAGGCTTTGATTTCAAAATGGCAAGAGCTCCACATGCTGATTTGAACAT 1499
OY 1711 TACAGTGTCCCGAGTCAACACAGAGAGAGAAAGAAATGGGCTTGAAGCTAGATCT 1770
Db 1500 TACAGTGTCCCGAGTCAACACAGAGAGAGAAAGAAATGGGCTTGAAGCTAGATCT 1559
OY 1771 CCTGCCATCAATCTGATGGGATTCACGCTGGAAGAGATGTATAGGCCCCAGCATGATC 1830
Db 1560 CCTGCCATCAATCTGATGGGATTCACGCTGGAAGAGATGTATAGGCCCCAGCATGATC 1619
OY 1831 CAAAGATCTTGAAGTCTCCAGAACCAACCATCATTTGAGAAATATCTATTTGTAACCT 1890
Db 1620 CAAAGATCTTGAAGTCTCCAGAACCAACCATCATTTGAGAAATATCTATTTGTAACCT 1679
OY 1891 GGGAGAAAGAACATGACATTTTGTCTCAGCTTCAGAAAACTTCAAGTGTCTCATACA 1950
Db 1680 GGGAGAAAGAACATGACATTTTGTCTCAGCTTCAGAAAACTTCAAGTGTCTCATACA 1739
OY 1951 GAAATTAACGCCACAGAGAGACAGAGTTAGAGATTGAAGAGCCCGGGCTGACCTCAT 2010
Db 1740 GAAATTAACGCCACAGAGAGACAGAGTTAGAGATTGAAGAGCCCGGGCTGACCTCAT 1799
OY 2011 GAGGTGGTTATGACATTTGAATGATATGCTTTGAAAGTACAGAGGAAATGGCAAGGAA 2070
Db 1800 GAGGTGGTTATGACATTTGAATGATATGCTTTGAAAGTACAGAGGAAATGGCAAGGAA 1859
OY 2071 AAGGACGAGGCGCTTGGGCTGCTGTAGACAGTGGACTATTCAGCAACAAAAAACCCAA 2130
Db 1860 AAGGACGAGGCGCTTGGGCTGCTGTAGACAGTGGACTATTCAGCAACAAAAAACCCAA 1919
OY 2131 GACGAATGAAGAAATATCATTTTCTGAATGTCTGTGCTTCCAACTCAAGACTT 2190
Db 1920 GACGAATGAAGAAATATCATTTTCTGAATGTCTGTGCTTCCAACTCAAGACTT 1979
OY 2191 CTGAATCAAAAGAACAGTTTGAAGAAATGGTTTGCAGGCTTCAAGGTGGAGAGAT 2250
Db 1980 CTGAATCAAAAGAACAGTTTGAAGAAATGGTTTGCAGGCTTCAAGGTGGAGAGAT 2039
OY 2251 GACAATGAGGCTCTTATGCTGCTCTTTCAAAGAAAGAAATATGAGAAAGAAACTG 2310
Db 2040 GACAATGAGGCTCTTATGCTGCTCTTTCAAAGAAAGAAATATGAGAAAGAAACTG 2099
OY 2311 CACAGGCAACCTGTGAGCCATAGGCTGTTTCAGCAAGTCCCATACAGTTCTGCAATGTG 2370
Db 2100 CACAGGCAACCTGTGAGCCATAGGCTGTTTCAGCAAGTCCCATACAGTTCTGCAATGTG 2159
OY 2371 GTATGAGAGTTGGCTTCAAGAAATGTCTGTGACACCTTGGATGCCAAATATAGGAGCT 2430
Db 2160 GTATGAGAGTTGGCTTCAAGAAATGTCTGTGACACCTTGGATGCCAAATATAGGAGCT 2219
OY 2431 GGCATATACCTTACCAAGAACCTCAAAAACTGGGAGAGAGAGCCAAAGAAATCTCTGCT 2490
Db 2220 GGCATATACCTTACCAAGAACCTCAAAAACTGGGAGAGAGAGCCAAAGAAATCTCTGCT 2279
OY 2491 GCAGATTAAGCTATGTATGTGTTGAGGCTGAAGTACTCACAGGCTCTTCTGCCAGGGA 2550

US-09-962-436-149/c
; Sequence 149, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Sopfel, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 149
; LENGTH: 284
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-149

Query Match 8.8%; Score 284; DB 10; Length 284;
Best Local Similarity 100.0%; Pred. No. 5,4e-64;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2951 GTGCTTATCTGTCTTATGCTTGGGGGTGGAGTAGATACCAATGAACACTT 3010
DB 284 GTGCTTATCTGTCTTATGCTTGGGGGTGGAGTAGATACCAATGAACACTT 225
QY 3011 TCAGGACCTTCCTCTCTTGCAGTTGTTCTTAATCTCCCTTACTAGAGAGATAATA 3070
DB 224 TCAGGACCTTCCTCTCTTGCAGTTGTTCTTAATCTCCCTTACTAGAGAGATAATA 165
QY 3071 TTTTCATATATGAAGAAATTTTCTAGTATATACGAGCCCTTTATTTCTAAAT 3130
DB 164 TTTTCATATATGAAGAAATTTTCTAGTATATACGAGCCCTTTATTTCTAAAT 105
QY 3131 GATGATAGTATAAATGTTAGATACAGATGATTTAGATTTCCAGAAATATTAT 3190
DB 104 GATGATAGTATAAATGTTAGATACAGATGATTTAGATTTCCAGAAATATTAT 45
QY 3191 AAGTGCCTTAGGTGAATAATCATCTTGTCTGATTAA 3234
DB 44 AAGTGCCTTAGGTGAATAATCATCTTGTCTGATTAA 1

RESULT 5
US-09-964-824A-308
; Sequence 308, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 308
; LENGTH: 260
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-308

Query Match 7.6%; Score 246.8; DB 10; Length 260;

Best Local Similarity 99.2%; Pred. No. 2.3e-54;
Matches 248; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2824 CTTACCTGGGGGACTAACCAATATGACCATGAGCGCCAAAGTGCTTGATA 2883
DB 1 CTTACCTGGGGGACTAACCAATATGACCATGAGCGCCAAAGTGCTTGATA 60
QY 2884 TATCCATGGGTTATCTATGAGTACTGAGGTTATGAAAGACTAGCCCATACTAG 2943
DB 61 TATCCATGGGTTATCTATGAGTACTGAGGTTATGAAAGACTAGCCCATACTAG 120
QY 2944 CATCTTAGGCTTTATCTGTTTATGCTTGGGGTGGGTAGTATACCAATGA 3003
DB 121 CATCTTAGGCTTTATCTGTTTATGCTTGGGGTGGGTAGTATACCAATGA 180
QY 3004 AACCTTCAGGACCTTCCTCTTGCAGTTGTTCTTAATCTCCCTTACTAGAGAG 3063
DB 181 AACCTTCAGGACCTTCCTCTTGCAGTTGTTCTTAATCTCCCTTACTAGAGAG 240
QY 3064 ATAAATATT 3073
DB 241 ATAAATATT 250

RESULT 6
US-09-796-692-7357
; Sequence 7357, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077 001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 7357
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7357

Query Match 7.0%; Score 226; DB 9; Length 238;
Best Local Similarity 99.6%; Pred. No. 5.3e-49;
Matches 237; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 01:59:24 ; Search time 2874 Seconds
(without alignments)
18274.866 Million cell updates/sec

Title: US-09-830-762-1

Perfect score: 3243
Sequence: 1 gggcttcgtctcctcgtggtg.....gtctgataaaaaaaaaa 3243

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_iny:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2052	63.3	2871	11	BC017463 Homo sapi
2	973	30.0	993	9	AL568673 AL568673
3	922	28.4	965	9	AL540976 AL540976
4	733.4	22.6	852	12	BG680130 BG680130
5	713.6	22.0	724	14	BQ014467 BQ014467
6	705.8	21.8	742	12	BG540093 BG540093

7	659.6	20.3	901	14	BQ923651	BQ923651	AGENCOURT
8	649.6	20.0	791	12 <td>BG284006</td> <td>BG284006</td> <td>602407302</td>	BG284006	BG284006	602407302
9	646.6	19.9	679	13 <td>BM664435</td> <td>BM664435</td> <td>UI-E-CL1</td>	BM664435	BM664435	UI-E-CL1
10	645	19.9	658	14 <td>BQ028103</td> <td>BQ028103</td> <td>UI-H-COO-</td>	BQ028103	BQ028103	UI-H-COO-
11	604.8	18.6	608	14 <td>BM829571</td> <td>BM829571</td> <td>K-EST0102</td>	BM829571	BM829571	K-EST0102
12	600.4	18.5	705	13 <td>B1857151</td> <td>B1857151</td> <td>603383747</td>	B1857151	B1857151	603383747
13	586.8	18.1	732	13 <td>AM851285</td> <td>AM851285</td> <td>IL3-CT022</td>	AM851285	AM851285	IL3-CT022
14	586.2	18.1	621	10 <td>B1836699</td> <td>B1836699</td> <td>603089634</td>	B1836699	B1836699	603089634
15	584.2	18.0	610	10 <td>AM851000</td> <td>AM851000</td> <td>IL3-CT022</td>	AM851000	AM851000	IL3-CT022
16	584	18.0	587	9 <td>AI307751</td> <td>AI307751</td> <td>tb24f11.x</td>	AI307751	AI307751	tb24f11.x
17	579.4	17.9	582	10 <td>AM850996</td> <td>AM850996</td> <td>IL3-CT022</td>	AM850996	AM850996	IL3-CT022
18	577.4	17.8	579	10 <td>AM851276</td> <td>AM851276</td> <td>IL3-CT022</td>	AM851276	AM851276	IL3-CT022
19	567.4	17.5	569	13 <td>BM503214</td> <td>BM503214</td> <td>1973b06.x</td>	BM503214	BM503214	1973b06.x
20	565.8	17.4	811	12 <td>BQ202125</td> <td>BQ202125</td> <td>RST21479</td>	BQ202125	BQ202125	RST21479
21	558.4	17.2	568	14 <td>BQ268851</td> <td>BQ268851</td> <td>1k21d03.x</td>	BQ268851	BQ268851	1k21d03.x
22	551.8	17.0	605	9 <td>AA151346</td> <td>AA151346</td> <td>z125b01.f</td>	AA151346	AA151346	z125b01.f
23	551	17.0	551	12 <td>BF111184</td> <td>BF111184</td> <td>7b4d12.x</td>	BF111184	BF111184	7b4d12.x
24	550	17.0	550	9 <td>AI860767</td> <td>AI860767</td> <td>w105d09.x</td>	AI860767	AI860767	w105d09.x
25	546	16.8	706	10 <td>BE612556</td> <td>BE612556</td> <td>601452069</td>	BE612556	BE612556	601452069
26	544.8	16.8	548	10 <td>BE326805</td> <td>BE326805</td> <td>hr65b06.x</td>	BE326805	BE326805	hr65b06.x
27	544.8	16.8	781	12 <td>BE962693</td> <td>BE962693</td> <td>601656069</td>	BE962693	BE962693	601656069
28	542	16.7	554	12 <td>BG768000</td> <td>BG768000</td> <td>602743759</td>	BG768000	BG768000	602743759
29	538	16.6	765	13 <td>B1855784</td> <td>B1855784</td> <td>603382722</td>	B1855784	B1855784	603382722
30	532.8	16.4	551	9 <td>AI670955</td> <td>AI670955</td> <td>wb12c08.x</td>	AI670955	AI670955	wb12c08.x
31	520	16.0	520	13 <td>BM503372</td> <td>BM503372</td> <td>1973b06.y</td>	BM503372	BM503372	1973b06.y
32	512	15.8	572	14 <td>BQ362709</td> <td>BQ362709</td> <td>MR2-ST022</td>	BQ362709	BQ362709	MR2-ST022
33	509.2	15.7	529	9 <td>AI670956</td> <td>AI670956</td> <td>wb12c09.x</td>	AI670956	AI670956	wb12c09.x
34	507	15.6	507	10 <td>AM474917</td> <td>AM474917</td> <td>xy21f12.x</td>	AM474917	AM474917	xy21f12.x
35	500	15.4	500	9 <td>AI129360</td> <td>AI129360</td> <td>q634a12.x</td>	AI129360	AI129360	q634a12.x
36	497.4	15.3	527	12 <td>BF057083</td> <td>BF057083</td> <td>7k15f01.x</td>	BF057083	BF057083	7k15f01.x
37	497.2	15.3	542	13 <td>B1010748</td> <td>B1010748</td> <td>MR2-EN009</td>	B1010748	B1010748	MR2-EN009
38	489	15.1	489	9 <td>AI304947</td> <td>AI304947</td> <td>qo20b12.x</td>	AI304947	AI304947	qo20b12.x
39	487	15.0	487	9 <td>AA431590</td> <td>AA431590</td> <td>zw70b08.f</td>	AA431590	AA431590	zw70b08.f
40	487	15.0	615	14 <td>BQ269118</td> <td>BQ269118</td> <td>1k21d03.y</td>	BQ269118	BQ269118	1k21d03.y
41	485	15.0	510	14 <td>NS4275</td> <td>NS4275</td> <td>yv66c11.sl</td>	NS4275	NS4275	yv66c11.sl
42	481	14.8	481	9 <td>AI653467</td> <td>AI653467</td> <td>tg94f02.x</td>	AI653467	AI653467	tg94f02.x
43	481	14.8	704	12 <td>BF684485</td> <td>BF684485</td> <td>602142974</td>	BF684485	BF684485	602142974
44	480.8	14.8	651	9 <td>AL602483</td> <td>AL602483</td> <td>DKP2p686P</td>	AL602483	AL602483	DKP2p686P
45	480	14.8	480	9 <td>AI523805</td> <td>AI523805</td> <td>tg96c06.x</td>	AI523805	AI523805	tg96c06.x

ALIGNMENTS

RESULT 1
BC017463
LOCUS
DEFINITION Homo sapiens, clone IMAGE:4862464, mRNA.
ACCESSION BC017463
VERSION BC017463.1 GI:16907195
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2871)
Strausberg,R.
Direct Submission
Submitted (13-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcapds-femail.nih.gov
Tissue Procurement: ATCC/DC/DNP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

REMARK
COMMENT

Leticia Hsiao, Martin Krzywinski, Reta Kulcsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mahewon, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvathi Saeehi, Jacqueline Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Matsija van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Maria.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 34 Row: h Column: 13
This clone has the following problem: no cloning site / microdeletion.

FEATURES

1. .2871

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4862464"
/tissue-type="Skin, melanotic melanoma, high MDR."
/clone_lib="NIH_MCC_49"

```

BASE COUNT	909 a	560 c	672 g	730 t
ORIGIN				

Query Match	63.3%	Score 2052;	DB 11;	Length 2871;
Best Local Similarity	95.2%;	Pred. NO. 0;		
Matches 2170; Conservative	0;	Mismatches 5;	Indels 105;	Gaps 1;

QY	136	GCACAAAGTGGGCGACCAATCTGGAAGAACTACAGTCTATGCTTTGAAGGCAAAAGGAAAT	195
Db	202	GAAAAAGTGGGCGACCAATCTATCTGGAAGCTACAGTCTATGCTTTGAAGGCAAAAGGAAAT	261
QY	196	AAACATTTAAAGACTCCCCGGGAGCTCGAGAGATGACCTTTCCATGSGTGGCGGAGCA	255
Db	262	AAACATTTAAAGACTCCCCGGGAGCTCGAGAGATGACCTTTCCATGSGTGGCGGAGCA	321
QY	256	GCAGCTTCAATGAAAAAAATCAGTGTAGATTAACCTCGCTCACCTTGTTCAGAAAGTC	315
Db	322	GCAGCTTCAATGAAAAA-----	339
QY	316	TTTGCTCAGATCTTTCCTCAGTGGAGAAAGGGAATACAGAAGAATGTCTCCCTACAAAG	375
Db	340	-----	339
QY	376	TGCTCAGAGACGTGGGCTCTGGAGAAAACTATAGTTGGCAATTCCTCATTAACCAAT	435
Db	340	---TCAGAGACTGGGCTCTGGAGAAAACTATAGTTGGCAATTCCTCATTAACCAAT	396
QY	436	GACTTCAAAATTTTAAAAAATATAGAGCGTCAGCTGTGTGAAGTCCCTCAGAAATTAATTT	495
Db	397	GACTTCAAAATTTTAAAAAATATATAGGCTCAGCTGTGTGAAGTCTCCAGAAATTAAGTTT	456
QY	496	GGCTGTATCTTACCCGTGTCTCTCAGTTCAAGAGCACAGCAAAATCTGTGCAAGTG	555
Db	457	GGCTGTATCTTACCCGTGTCTCTCAGTTCAAGAGCACAGCAAAATCTGTGCAAGTG	516
QY	556	TTTCGAAAAATGCTGACCTCCTAGAGTAAAGTATACGTCGTGGAAGATGACCTCACCA	615
Db	517	TTTCGAAAAATGCTGACCTCCTAGAGTAAAGTATATAGTCTGGAAGATGACCTCACCA	576
QY	616	CATGCTGTGATGCTGTGTGTAATCAGACCAATGAAGATCTTCTCATGSGGAGGAGCTG	675
Db	577	CATGCTGTGATGCTGTGTGTAATCAGACCAATGAAGATCTTCTCATGSGGAGGAGCTG	636
QY	676	GCCCTGGCCCTGGTAAAAAGCTGGTGGATTTGAAATCCAGAAAGAGAGCAACAGTTTGTT	735
Db	637	GCCCTGGCCCTGGTAAAAAGCTGGTGGATTTGAAATCCAGAAAGAGAGCAACAGTTTGTT	696
QY	736	GCACAGATATGCTAAAGTCTCAGCTGCTGAGATAGCTGTCAACGGAGACAGGAGGCTTCCC	795
Db	697	GCACAGATATGCTAAAGTCTCAGCTGCTGAGATAGCTGTCTCAACGGAGACAGGAGGCTTCCC	756
QY	796	TGCAAACAGATCATCATGCTGTGGGCTCGGTGGATGGAATGGAGTAAACAGGGAATGT	855

Db	757	TGCAACAGATCATTCATGCTGTGTGGGCGCTCGGTGAGATGGAATGGGATTAACAGGGATGT	816
Qy	856	ACTGGAAGCTGCAGAGGGCCATGTGAATTTCTGAATTAATGTCATCTATAAAAAATACT	915
Db	817	ACTGGAAGCTGCAGAGGGCCATTTAGATTAATTCGATTTATGCATCTATAAAAAATACT	876
Qy	916	CACATTAAGCACTACCATTTCACGCTTGAGCTGAGCTCTGGGATTTTTCAGTTCCCGCTGAAT	975
Db	877	CACATTAAGCACTACCATTTCACGCTTGAGCTCTGGGATTTTTCAGTTCCCGCTGAAT	936
Qy	976	TTTGTCACAAGACTATTGTAGAGACTATCCGGGTTAGTTTGCAAGGGAAGCCATGATG	1035
Db	937	TTTGTCACAAGACTATTGTAGAGACTATCCGGGTTAGTTTGCAAGGGAAGCCATGATG	996
Qy	1036	AGTAAATTGAAAGAAATTCACCTGGTGAGCAATGAGACCTACTGTTGCTGCCCTTAAA	1095
Db	997	AGTAAATTGAAAGAAATTCACCTGGTGAGCAATGAGACCTACTGTTGCTGCCCTTAAA	1056
Qy	1096	GCTGCTTCACAATTCATCCTAGGGAAGATGAGCTGGGACAAAGAAACCACCCCTCTTTC	1155
Db	1057	GCTGCTTCACAATTCATCCTAGGGAAGATGAGCTGGGACAAAGAAACCACCCCTCTTTC	1116
Qy	1156	AATGCAATGGTGTGGAACAACCGACCCCTCGATTTGTCAGGGCCCATTTGAATGGCAG	1215
Db	1117	AATGCAATGGTGTGGAACAACCGACCCCTCGATTTGTCAGGGCCCATTTGAATGGCAG	1176
Qy	1216	ACGGCAGATTAATTTGTTAAATTCGTAAACCACACATGATATTACAGTTGGAACCTGTGGCA	1275
Db	1177	ACGGCAGATTAATTTGTTAAATTCGTAAACCACACATGATATTACAGTTGGAACCTGTGGCA	1236
Qy	1276	AAGTCAATTTCTACACACAGAGAGTTGGAATGAATGGAAATTTCTTGCCACAAAGGCT	1335
Db	1237	AAGTCAATTTCTACACACAGAGAGTTGGAATGAATGGAAATTTCTTGCCACAAAGGCT	1296
Qy	1336	AAACAGTTTAACAGGCGCCAGTGGTGAACGTGACACAAAAGATTTAACTGTTGCTGTAA	1395
Db	1297	AAACAGTTTAACAGGCGCGCCAGTGGTGAACGTGACACAAAAGATTTAACTGTTGCTGTAA	1356
Qy	1396	TATATATACCATGTACTGTGGCATTCAGAAATTTCCATAACCTCAGATATTAAAACATGCA	1455
Db	1357	TATATATACCATGTACTGTGGCATTCAGAAATTTCCATAACCTCAGATATTAAAACATGCA	1416
Qy	1456	ATGAGAGAGTGTGGGAAAAATGCAATTTGAGCAAAATATTAACCTGCATTTCTTCCCTGCC	1515
Db	1417	ATGAGAGAGTGTGGGAAAAATGCAATTTGAGCAAAATATTAACCTGCATTTCTTCCCTGCC	1476
Qy	1516	CTTGGGAGCTGGAAACATGGAAATTAAGAAGGAACACGACGAGATTTGTTTGATGAA	1575
Db	1477	CTTGGGAGCTGGAAACATGGAAATTAAGAAGGAACACGACGAGATTTGTTTGATGAA	1536
Qy	1576	GTTTAAACATTTGCCCAAAGACCATGTAAACACACAGTTAACTGTAAAAATTTTGATCTTT	1635
Db	1537	GTTTAAACATTTGCCCAAAGACCATGTAAACACACAGTTAACTGTAAAAATTTTGATCTTT	1596
Qy	1636	CCAACAGATTTGGAGATATATAAGGCTTTCACTGTGAATGGCAAAAGGCTCCAGATG	1695
Db	1597	CCAACAGATTTGGAGATATATAAGGCTTTCACTGTGAATGGCAAAAGGCTCCAGATG	1656
Qy	1696	CTGAGTTTGAACAATTAACAGTGTGCCCCAGTCAACCAAGAGAGGAAAAAGGAAAAATGGG	1755
Db	1657	CTGAGTTTGAACAATTAACAGTGTGCCCCAGTCAACCAAGAGAGGAAAAAGGAAAAATGGG	1716
Qy	1756	CTTGAACCTTGATCTCTCGTCATCATCTGATGGGATTTCAACGTGGAAGAGATATAG	1815
Db	1717	CTTGAACCTTGATCTCTCGTCATCATCTGATGGGATTTCAACGTGGAAGAGATATAG	1776
Qy	1816	GCCACACCATGATCCAAAGAAATCCTGATCTTCCAGAACCAACCATCATTTGAGAAATAT	1875
Db	1777	GCCACACCATGATCCAAAGAAATCCTGATCTTCCAGAACCAACCATCATTTGAGAAATAT	1836
Qy	1876	CATATTCTGACTTTGGGAGAAAGAAATGACATTTTGTCTACAGCTTCAGAAAACTTCA	1935

Db	Accession	Gene	Species	Strand	Length	Start	End	Score	Expect	Method	Program	Version	Database	Source	Comment
Db	1837	CATATTCGTACCTTTGGAGAAAGACATACATATTTGTCTCAGCTTCAAGAAAACCTTCA	human	+	1896	1	1896	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
QY	1936	AGTGTCTCCATCAGCAAAATTTATCAGCCACAGAAAGACAGATTAGACATTTGAAGAGACC	human	+	1995	1	1995	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
Db	1897	AGTGTCTCCATCAGCAAAATTTATCAGCCACAGAAAGACAGATTAGACATTTGAAGAGACC	human	+	1996	1	1996	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
QY	1996	CGGGCTGACCTTCATTGAGGCTGTTATGACACTTGAAGATATGCTTTTGTAAAGTACAGAG	human	+	2055	1	2055	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
Db	1957	CGGGCTGACCTTCATTGAGGCTGTTATGACACTTGAAGATATGCTTTTGTAAAGTACAGAG	human	+	2016	1	2016	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
QY	2056	GAATGGCAAGAAAAAGACGACGAGCCCTTTGGGCGCTCGTAGACACAGTGCATATTCAG	human	+	2115	1	2115	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
Db	2017	GAATGGCAAGAAAAAGACGACGAGCCCTTTGGGCGCTCGTAGACACAGTGCATATTCAG	human	+	2076	1	2076	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
QY	2116	CAACAAAAAACCCAAAGACGAATGAAGAAATATCATATTTCTGAATGTCTCGTGCGCT	human	+	2175	1	2175	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
Db	2077	CAACAAAAAACCCAAAGACGAATGAAGAAATATCATATTTCTGAATGTCTCGTGCGCT	human	+	2136	1	2136	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
QY	2176	CCAACTCAAGAGCTTCTAGATCAAAAAGAACAGTTTGAATAATGTGTTTCAGAGTTCTA	human	+	2235	1	2235	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
Db	2137	CCAACTCAAGAGCTTCTAGATCAAAAAGAACAGTTTGAATAATGTGTTTCAGAGTTCTA	human	+	2196	1	2196	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
QY	2236	AAGGTGGAGAGATGACAAATGAGAGTCCCTTATGTGCTGCTTCAAAAGAAAGAAATG	human	+	2295	1	2295	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
Db	2197	AAGGTGGAGAGATGACAAATGAGAGTCCCTTATGTGCTGCTTCAAAAGAAAGAAATG	human	+	2256	1	2256	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
QY	2296	ATGGAAGAAAAACCTGCACAGCAACCTTGAGCCATAGAGCTGTTTCACAGTCCCATAC	human	+	2355	1	2355	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
Db	2257	ATGGAAGAAAAACCTGCACAGCAACCTTGAGCCATAGAGCTGTTTCACAGTCCCATAC	human	+	2316	1	2316	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
QY	2356	CAGTCTGCAATGTGTATGACAGAGTTGGCTTTCAAGAAATGTACTGCACACCTTGCGAT	human	+	2415	1	2415	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	
Db	2317	CAGTCTGCAATGTGTATGACAGAGTTGGCTTTCAAGAAATGTACTGCACACCTTGCGAT	human	+	2376	1	2376	100.0	1e-100	BLAST	NCBI	1.0	NCBI	human	

Query Match	30.0%;	Score 973;	DB 9;	Length 993;
Best Local Similarity	98.8%;	Pred. No. 1.7e-242;		
Matches 981;	Conservative	8;	Mismatches 3;	Indels 1;
				Gaps 1;
QY 2206 CAGTTTGAAAATGTTGGTTTGCAGGTTCTTAAAGTGTGAGAGATGAGACATGAGGTCCTT	:			2265
DB 993 CAGTTTAAAAATGTGTGGTGGCAGGTTCTAAAGTGTGAGAGATGAGACATGAGGTCCTT	:			934
QY 2266 AAGGCTGCTTTCAAAAGAAAGAA -GAAAAATGTGAGAGAAAAACATGACAGCAACCTGT	:			2324
DB 933 AAGGCTKCTTTTTCAAAAGAAAGAAADAAAAATGATGAGAGAAAAACATGACAGCAACCTGT	:			874
QY 2325 GAGCCATAGGCTGTTTTCACAGTCCCATATACAGTTCTCCATGTGGATGTGATGAGAGTTGG	:			2384
DB 873 GAGCCATAGGTGTGTTTTCACAGCAAGTCCCATATACAGTTCTCCATGTGGATGTGAGAGTTGG	:			814
QY 2385 CTTTCAAGAAATGTACTCGACACCTTGGCATCCAAAATACGAGAGCTGGCATATACCTTAC	:			2444
DB 813 CTTTCAAGAAATGTACTCGACACCTTGGCATCCAAAATACGAGAGCTGGCATATACCTTAC	:			754
QY 2445 CAAAGAACCTTCAAAAACCTGTGGCAGAGAGAGCCAAAGAAATCTGCTGCAGATTAACCTGAT	:			2504
DB 753 CAAAGAACCTTCAAAAACCTGTGGCAGAGAGAGCCAAAGAAATCTGCTGCAGATTAACCTGAT	:			694
QY 2505 CATATGTGTTGAGGCTGAAGTACTGACAGGCTTCTTCCAGAGGACATCCGTTAAATAT	:			2564
DB 693 CATATGTGTTGAGGCTGAAGTACTGACAGGCTTCTTCCAGAGGACATCCGTTAAATAT	:			634
QY 2565 TGTTCCTCCACACCATGTAGTCTGTGAGCTATAGATGGTCAATGACAGTGTGGTACAAATGT	:			2624
DB 633 TGTTCCTCCACACCATGTAGTCTGTGAGCTATAGATGGTCAATGACAGTGTGGTACAAATGT	:			574
QY 2625 CTCACAGCCCTGAACCTTTGTATTTTATAGTGGCATGCAGGCTATATACCTCAATATTGG	:			2684
DB 573 CTCACAGCCCTGAACCTTTGTATTTTATAGTGGCATGCAGGCTATATACCTCAATATTGG	:			514
QY 2685 GACATGACACCCAGGAATATGTACAGTACACAAGATTAATCATATAGAGCCAAATGAGACCCCT	:			2744
DB 513 GACATGACACCCAGGAATATGTACAGTACACAAGATTAATCATATAGAGCCAAATGAGACCCCT	:			454
QY 2745 TGCACAGCATCCTTGGAGGGGATTGCGCAAGTGGCAGCCCTGTTGATTAATCTCTACATCA	:			2804
DB 453 TGCACAGCATCCTTGGAGGGGATTGCGCAAGTGGCAGCCCTGTTGATTAATCTCTACATCA	:			394
QY 2805 TTTTAAACAGCTGTATGGCTTATCCTTGGGTGAACTAACCAATATATGACATCGATGGC	:			2864
DB 393 TTTTAAACAGCTGTATGGCTTATCCTTGGGTGAACTAACCAATATATGACATCGATGGC	:			334
QY 2865 TCAAAAGAGTGGCTTGAATATATCCCATGGGTTATCTGTATGACATGACATGGGTTATTGAA	:			2924
DB 333 TCAAAAGAGTGGCTTGAATATATCCCATGGGTTATCTGTATGACATGACATGGGTTATTGAA	:			274
QY 2925 AGGACTAGCCACATATACATCTAGTACCTTATCCTTATCTGTATATGTCCTTGGGTTGGG	:			2984
DB 273 AGGACTAGCCACATATCTAGTACCTTATCCTTATCCTTATCTGTATATGTCCTTGGGTTGGG	:			214
QY 2985 GTTAGGTAGTATACCAATGAAACATTTTACAGACCTTCTCTCTGTGACAGTTGTCTTTA	:			3044
DB 213 GTTAGGTAGTATACCAATGAAACATTTTACAGACCTTCTCTCTGTGACAGTTGTCTTTA	:			154
QY 3045 ATCTCCTTACTAGAGAGAAATTAATATTTTGGATATATGAAGAAATTTTCTAGATATAT	:			3104
DB 153 ATCTCCTTACTAGAGAGAAATTAATATTTTGGATATATGAAGAAATTTTCTAGATATAT	:			94
QY 3105 AACGCGAGGCTTTTATTTTCTAAATATGATAGATATAAAATATTTAGGATTAACGAATG	:			3164
DB 93 AACGCGAGGCTTTTATTTTCTAAATATGATAGATATAAAATATTTAGGATTAACGAATG	:			34
QY 3165 ATTTTATGATTTTCCAGAGAAATATTTAAAGTCG 3197	:			
DB 33 ATTTTATGATTTTCCAGAGAAATATTTTAAARTGC 1	:			

```

RESULT 3
AL540976          965 bp      mRNA      linear      EST 16-FEB-2001
LOCUS             AL540976.LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE004Y014 5 prime
DEFINITION        , mRNA sequence.
ACCESSION         AL540976.1 GI:12871606
VERSION           EST.
KEYWORDS          human.
SOURCE            Homo sapiens
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         1 (bases 1 to 965)
AUTHORS           Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE             Full-length cDNA libraries and normalization
JOURNAL           Unpublished (2001)
COMMENT           Contact: Genoscope
                  Genoscope - Centre National de Sequencage
                  BP 191 91006 Evry cedex - France
                  Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
  source
    1..965
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="CS0DE004Y014"
      /clone_1ib="LTI_FL002_PL1"
      /lab_host="DH10B"
      /note="organ: placenta; Vector: pCMVSPORT 6; 1st strand
      cDNA was primed with a NotI-190(47) primer. Five prime
      end enriched, double-stranded cDNA was digested with Not I
      and cloned into the Not I and Eco RV sites of the
      pCMVSPORT 6 vector. Library was constructed by Life
      Technologies. Contact : Feng Liang Life Technologies, a
      division of Invitrogen 9800 Medical Center Drive Rockville
      , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
      fliang@lifetech.com URL :
      http://fulllength.invitrogen.com"

BASE COUNT      280 a      194 c      243 g      247 t      1 others
ORIGIN
Query Match      28.4%; Score 922; DB 9; Length 965;
Best Local Similarity 99.5%; Pred. No. 3.3e-229;
Matches 955; Conservative 1; Mismatches 1; Indels 3; Gaps 3;

QY 137 CAGAACTGGGCCACCATCTGGAACCTACAGTCTATGCTTGAAGCGAAAGGGAATA 196
    1 |
DB 8 CGAACTGGGCCACCATCTGGAACCTACAGTCTATGCTTGAAGCGAAAGGGAATA 67

QY 197 AACATTTAAGACTCCCGCGGACCTGAGAGATGAGCTTTTCATGATGCGCGAGCAG 256
    1 |
DB 68 AACATTTAAGACTCCCGCGGACCTGAGAGATGAGCTTTTCATGATGCGCGAGCAG 127

QY 257 CAGCTTACATGAAATATGAGTAGATTACTCGCTCTCACTCTTGTTCAGAAAGTCT 316
    1 |
DB 128 CAGCTTACATGAAATATGAGTAGATTACTCGCTCTCACTCTTGTTCAGAAAGTCT 187

QY 317 TTGGCTCAGATCTTCTCTCAGAGAGAAAGGGAATACAGAAAGTCTCCCTACAAGT 376
    1 |
DB 188 TTGGCTCAGATCTTCTCTCAGAGAGAAAGGGAATACAGAAAGTCTCCCTACAAGT 247

QY 377 GCTCAGAGACTGCTCTTTGGAGAAATATAGTGGCAAAATCCCATTAACCAAGT 436
    1 |
DB 248 GCTCAGAGACTGCTCTTTGGAGAAATATAGTGGCAAAATCCCATTAACCAAGT 307

QY 437 ACTTCAAAATTTTAAAAATTAATGAGCTGACCTGTGAAGTCTTCAGAAATAGTTTG 496
    1 |
DB 308 ACTTCAAAATTTTAAAAATTAATGAGCTGCA-CTGTGTAAGTCTTCAGAAATAGTTTG 366

QY 497 GCTGTATCTCTACCCGCGTCTCTCCAGTTCAGAGAAAGCAACCAATCTCTGCAAGTGT 556
    1 |
DB 367 GCTGTATCTCTACCCGCGTCTCTCCAGTTCAGAGAAAGCAACCAATCTCTGCAAGTGT 426

```

```

QY 557 TCAGAAAAATGCTGACTCCTAGAGATAGATTACACTGTGAAAGATGACCTCACACAC 616
    1 |
DB 427 TCAGAAAAATGCTGACTCCTAGAGATAGATTACAGTGTGGAAGATGACCTCACACAC 486

QY 617 ATGCTGTATCTCTGCTGTGGAATGAGCAATGAGAAATCTTCGATGGGGAGGCTGTG 676
    1 |
DB 487 ATGCTGTATCTCTGCTGTGGAATGAGCAATGAGAAATCTTCGATGGGGAGGCTGTG 546

QY 677 CCTTGCCCTGTGTAAGCTGTGATTTGAATCCAAAGAGCAACAGTTTGTG 736
    1 |
DB 547 CCTTGG-CCTGTTAAAGCTGTGATTTGAATCCAAAGAGCAACAGTTTGTG 605

QY 737 CCAGATATGTTAAAGTGTGACCTGTGATGATGCTGTCAAGGAGAGGAGGCTTCCCT 796
    1 |
DB 606 CCAGATATGTTAAAGTGTGACCTGTGATGATGCTGTCAAGGAGAGGAGGCTTCCCT 665

QY 797 GCAACAGATCATTCATGCTGTGTTGGCCCGGCGGTGATGGAATGGGTAACAGGGATGTA 856
    1 |
DB 666 GCAACAGATCATTCATGCTGTGTTGGCCCGGCGGTGATGGAATGGGTAACAGGGATGTA 725

QY 857 CTGAAAGCTGAGAGAGGCGCATTTGTAATTTCTGAATTTGATCTATTAATAATCTC 916
    1 |
DB 726 CTGAAAGCTGAGAGAGGCGCATTTGTAATTTCTGAATTTGATCTATTAATAATCTC 785

QY 917 ACATTAAGACATGACATTCACGCTTGAGCTTGGGATTTTTCAGTTCCCTCTGAATT 976
    1 |
DB 786 ACATTAAGACATGACATTCACGCTTGAGCTTGGGATTTTTCAGTTCCCTCTGAATT 845

QY 977 TGTGTATC-AAAGACTATTTAGAGACTATCCGGGTAGTTTCAAGGGAGCAATGANG 1035
    1 |
DB 846 TGTGTATC-AAAGACTATTTAGAGACTATCCGGGTAGTTTCAAGGGAGCAATGANG 905

QY 1036 ACTAATTTGAAGAAATTTACCTGTGTGAGCAATGAGAGACCTACTGTTCTCTTTANA 1095
    1 |
DB 906 ACTAATTTGAAGAAATTTACCTGTGTGAGCAATGAGAGACCTACTGTTCTCTCTTTANA 965

RESULT 4
BG680130          852 bp      mRNA      linear      EST 01-MAY-2001
LOCUS             BG680130
DEFINITION        602628288F1 NCI_CGAP_Skn4 Homo sapiens cDNA IMAGE:4753386 5',
    mRNA sequence.
ACCESSION         BG680130
VERSION           BG680130.1 GI:13911527
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         1 (bases 1 to 852)
AUTHORS           NIH-MGC http://mgc.ncl.nih.gov/.
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cga@bs-remail.nih.gov
                  Tissue Procurement: James Cleaver, M.D.
                  cDNA Library Preparation: Life Technologies, Inc.
                  DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM10612 row: 0 column: 21
                  High quality sequence stop: 792.
                  Location/Qualifiers
                    1..852
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4753386"
                      /clone_1ib="NCI_CGAP_Skn4"
                      /tissue_type="squamous cell carcinoma"
                      /lab_host="DH10B (TI phage-resistant)"

```

/note="Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT 275 a 177 c 204 g 196 t

ORIGIN

Query Match 22.6%; Score 733.4; DB 12; Length 852;
Best Local Similarity 97.0%; Pred. No. 5.3e-180;
Matches 822; Conservative 0; Mismatches 16; Indels 9; Gaps 7;

QY 1786 ATGGGATTCAACGTGGAAGAGATGTATGAGGCCCAAGCATGATCCAAAGATCCYGAAT 1845
|||||
Db 1 ATGGGATTCAACGTGGAAGAGATGTATGAGGCCCAAGCATGATCCAAAGATCCYGAAT 60

QY 1846 CTCGGAACCCACCCATCATGATGAGAAATATCATATCTGTACTCTGGGAGAAAGAACAT 1905
|||||
Db 61 CTCGGAACCCACCCATCATGATGAGAAATATCATATCTGTACTCTGGGAGAAAGAACAT 120

QY 1906 GACATTTTGTCTCAGCTTCAGAAAACCTTCAAGTGTCTCCATCAGACAGAAATATACGCCCA 1965
|||||
Db 121 GACATTTTGTCTCAGCTTCAGAAAACCTTCAAGTGTCTCCATCAGACAGAAATATACGCCCA 180

QY 1966 GGAAGCAGAGATTAGAGATTGAAGAGCCCGGGCTGACCTCATTTAGAGTGGTTATGAAC 2025
|||||
Db 181 GGAAGCAGAGATTAGAGATTGAAGAGCCCGGGCTGACCTCATTTAGAGTGGTTATGAAC 240

QY 2026 ATTGAAGATATGCTTTGTTAAATACAGAGAGAAATGCGAAGAAAAGAGAGAGCCCTT 2085
|||||
Db 241 ATTGAAGATATGCTTTGTTAAATACAGAGAGAAATGCGAAGAAAAGAGAGAGCCCTT 300

QY 2086 TGGCCCTGTTAGAGACACTGACCTATTTCAGCAACAAAAACCCAGACGAATGAAGAA 2145
|||||
Db 301 TGGCCCTGTTAGAGACACTGACCTATTTCAGCAACAAAAACCCAGACGAATGAAGAA 360

QY 2146 AATATCATATTTCTGAAAATGTCCTGCTCCACACTCAAGAGCTTCTAGATCAAAAAGAA 2205
|||||
Db 361 AATATCATATTTCTGAAAATGTCCTGCTCCACACTCAAGAGCTTCTAGATCAAAAAGAA 420

QY 2206 CAGTTTGAAA-AATGSGTTTGCAGGTTCTTA-AGTGGAGAAAGATAGACATGAGGTC 2263
|||||
Db 421 CAGTTTGAAA-AATGSGTTTGCAGGTTCTTA-AGTGGAGAAAGATAGACATGAGGTC 480

QY 2264 TTATGGCTGCTTTCAGAAAGAAAGAAATGATGGAAGAAAACCTGCACAGCAACCTG 2323
|||||
Db 481 TTATGGCTGCTTTCAGAAAGAAAGAAATGATGGAAGAAAACCTGCACAGCAACCTG 540

QY 2324 TGAGCCATAGGCTGTTTACAGCAAGTCCCATACAGTTCGCAATGTGTTAGCAGAGTTG 2383
|||||
Db 541 TGAGCCATAGGCTGTTTACAGCAAGTCCCATACAGTTCGCAATGTGTTAGCAGAGTTG 600

QY 2384 GCTTTC-AAGAATGTACGCGACACTTGGCGATCCAAAATAGGAGCTGATATACCTT 2442
|||||
Db 601 GCTTTCACAGATGTACTCGACACCTTGGCGATCCAAAATAGGAGCTGATATACCTT 660

QY 2443 ACCAAGAACCCTCAAAACCTGGCGAG-AAGGCCAAGAAATCTGTGCT--GCAGATAG 2499
|||||
Db 661 ACCAAGAACCCTCAAAACCTGGCGAGAGGCCAAGAAATCTGTGCTTGCAGATTAAG 720

QY 2500 CTGATCTATGTGTTT-GAGGCTGAAGTACTCAGAGCTTCTTCTGCCAGGACATCCGTT 2558
|||||
Db 721 CTGATCTATGTGTTT-GAGGCTGAAGTACTCAGAGCTTCTTCTGCCAGGACATCCGTT 780

QY 2559 AAATTTGTGTCGCCACAGAGTCTGGAGCTATA--GATGTCATGACAGTGTGTT 2616
|||||
Db 781 AAATTTGTGTCGCCACAGAGTCTGGAGCTATA--GATGTCATGACAGTGTGTT 840

QY 2617 GACAAATG 2623
|||||
Db 841 GACCAAG 847

RESULT 5

B0014467/c
LOCUS B0014467 724 bp mRNA linear EST 26-MAR-2002
DEFINITION U-H-ED1-axt-k-07-0-UI.s1 NCI-CGAP_ED1 Homo sapiens CDNA clone
IMAGE:5833470 3', mRNA sequence.

ACCESSION B0014467
VERSION B0014467.1 GI:19739368
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 724)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-retail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLY-A=yes.

FEATURES
source
location/Qualifiers
1..724
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5833470"
/clone_lib="NCI CGAP ED1"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Left Public Bone; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI CGAP ED1 is a normalized CDNA library containing the following tissue(s): Chondrosarcoma cell line C5. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand CDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded CDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand CDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGCTT.
TAG_LIB=UI-H-ED1
TAG_TISSUE=chondrosarcoma
TAG_SEQ=CGTCAAGCTT"

BASE COUNT 227 a 145 c 135 g 216 t 1 others

ORIGIN

Query Match 22.0%; Score 713.6; DB 14; Length 724;
Best Local Similarity 99.3%; Pred. No. 7.3e-175;
Matches 716; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2533 ACTACTCAGGCTTCTTCTGCGAGGAGACATCGTTAAATATGTTCCCCACACTGAG 2582
|||||
Db 724 AGTACTCAGGCTTCTTCTGCGAGGAGACATCGTTAAATATGTTCCCCACACTGAG 665

QY 2583 TCCTGAGACTAATAGATGTCATGACAGAGTGTGTTACAAATGTCCTCAGACCCGGAACCTT 2642
|||||
Db 664 TCCTGAGACTAATAGATGTCATGACAGAGTGTGTTACAAATGTCCTCAGACCCGGAACCTT 605

QY 2643 TGTATATTTTATGTCATGACAGGCTATACCTCAGATTTTGGACATCACCACAGAAATA 2702
|||||
Db 604 TGTATATTTTATGTCATGACAGGCTATACCTCAGATTTTGGACATCACCACAGAAATA 545

QY 2703 TGTACAGTCACAAAGATTACTATCAGAGCAATGAGACCTTTGACAGCATCTCTGGAG 2762
|||||

Db	544	TGTACAGTCACAAGATTACTCATCAAGGACCATGAGACCCTTTGGACAGCATTCTTGGAG	485
Oy	2763	GCGATTCCGAAGTGGCAGGCCCTTGATTAACTCTCTACATCTTTTAACAGCTGGTATGG	2822
Db	484	GGGATTCCCAAATGGCAGCCCTGTGATTAACTCTACATCTTTTAACAGCTGGTATGG	425
Oy	2823	CCTTACCTTGGGTGAAGTAACCAAAATAATGACCATTCGATGGCTCAAGAAGTGGCTTGAA	2882
Db	424	CTTTACCTTGGGTGAAGTAACCAAAATAATGACCATTCGATGGCTCAAGAAGTGGCTTGAA	365
Oy	2883	ATATCCATGGGTTATCTGTATGACCTGAGCTGGTTATTTGAAAAGACTAGCCATACTA	2942
Db	364	ATATCCATGGGTTATCTGTATGACCTGAGCTGGTTATTTGAAAAGACTAGCCATACTA	305
Oy	2943	GCATTTAGTGGCTTATCTGTCTTATGCTCTTGGGGTGGGGTAGTATACCAAAAG	3002
Db	304	GCATTTAGTGGCTTATCTGTCTTATGCTCTTGGGGTGGGGTAGTATACCAAAAG	245
Oy	3003	AAACACTTTCAGAGACCTTCCTCTCTCTGCACTGTTCTTTATCTCCTTACTAGAGA	3062
Db	244	AAACACTTTCAGAGACCTTCCTCTCTGCACTGTTCTTTATCTCCTTACTAGAGA	185
Oy	3063	GATAAATATTTTGCATPATATGACAGAAATTTTGTAGTATATPAACGAGCCTTTTATT	3122
Db	184	GATAAATATTTTGCATPATATGACAGAAATTTTGTAGTATATPAACGAGCCTTTTATT	125
Oy	3123	TCTAAATGATGATGTGTTTTAAAAAGCTTACAGATTAACAGAAATATTTAGATTTCCAGAG	3182
Db	124	TCTAAATGATGATGTGTTTTAAAAAGCTTACAGATTAACAGAAATATTTAGATTTTCCAGAG	65
Oy	3183	AATATTTATTAAGCTGTTAGTGATGAAAAATAATCATCTTTGTCTGATTTAAAAA	3242
Db	64	AATATTTATTAAGCTGTTAGTGATGAAAAATAATCATCTTTGTCTGAAAAA	5
Oy	3243	A 3243	
Db	4	A 4	
RESULT 6	BGS40093	742 bp mRNA Linear EST 03-APR-2001	
LOCUS	602586929P1 NIH_MGC_77 Homo sapiens CDNA IMAGE:4693350 5'		
DEFINITION	mRNA sequence.		
ACCESSION	BGS40093		
VERSION	BGS40093.1 GI:13532326		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strusberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: CLONTECH Laboratories, Inc. cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNCMI515 row: 3 column: 07 High quality sequence stop: 724. Location/Qualifiers 1..742		
FEATURES	source		

Query Match	Best Local Similarity	Score	DB	Length	742;
Matches	721, Conservative	0;	Mismatches	8;	Indels 1; Gaps 1;
QY 2357	AGTTGTGCAATGTGTATGACAGATGTGGCTTTC	21.8%;	202 a	158 c	162 g 219 t 1 others
Db 1	AGTTGTGCAATGTGTATGACAGATGTGGCTTTC	98.8%;	0;	0;	0;
QY 2417	CAAAATACGAGAGCTGCATATCTACTTCACCA	705.8;	DB 12;	Length 742;	
Db 61	CAAAATACGAGAGCTGCATATCTACTTCACCA	705.8;	DB 12;	Length 742;	
QY 2477	AGAAATCTCTGCTGCAGATAAGCTGATCTATG	705.8;	DB 12;	Length 742;	
Db 121	AGAAATCTCTGCTGCAGATAAGCTGATCTATG	705.8;	DB 12;	Length 742;	
QY 2537	TCTTGTGCGAGGAGCATCCGTTAAATATGTGT	705.8;	DB 12;	Length 742;	
Db 181	TCTTGTGCGAGGAGCATCCGTTAAATATGTGT	705.8;	DB 12;	Length 742;	
QY 2597	ATGTCATGACAGTGTGGTTGACATGTCTCCAG	705.8;	DB 12;	Length 742;	
Db 240	ATGTCATGACAGTGTGGTTGACATGTCTCCAG	705.8;	DB 12;	Length 742;	
QY 2657	GCATGACAGGCTATACCTCAGTATTTGTGACA	705.8;	DB 12;	Length 742;	
Db 300	GCATGACAGGCTATACCTCAGTATTTGTGACA	705.8;	DB 12;	Length 742;	
QY 2717	ATTACTCATCAGGACCAATGAGACCTTTCGACA	705.8;	DB 12;	Length 742;	
Db 360	ATTACTCATCAGGACCAATGAGACCTTTCGACA	705.8;	DB 12;	Length 742;	
QY 2777	GCAGCCCTGTTGATTAACTCTACATCATTTAA	705.8;	DB 12;	Length 742;	
Db 420	GCAGCCCTGTTGATTAACTCTACATCATTTAA	705.8;	DB 12;	Length 742;	
QY 2837	AACTAACCAAAATATGACCATTGATGCGTCAAA	705.8;	DB 12;	Length 742;	
Db 480	AACTAACCAAAATATGACCATTGATGCGTCAAA	705.8;	DB 12;	Length 742;	
QY 2897	ATCTGTATGAGATGACTGGGTATTGAAAGAGCT	705.8;	DB 12;	Length 742;	
Db 540	ATCTGTATGAGATGACTGGGTATTGAAAGAGCT	705.8;	DB 12;	Length 742;	
QY 2957	TTATGTGCTTTATGCTCTTGGGGTGGGGTAA	705.8;	DB 12;	Length 742;	
Db 600	TTATGTGCTTTATGCTCTTGGGGTGGGGTAA	705.8;	DB 12;	Length 742;	
QY 3017	CCCTTCCTTCCTTCGAGTGTGCTTTTAATCTC	705.8;	DB 12;	Length 742;	
Db 660	CCCTTCCTTCCTTCGAGTGTGCTTTTAATCTC	705.8;	DB 12;	Length 742;	
QY 3077	ATATTAATGAA 3086	705.8;	DB 12;	Length 742;	
Db 720	ATATTAATGAA 729	705.8;	DB 12;	Length 742;	

```

DEFINITION AGENCOURT_8854013 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6473545
ACCESSION 5', mRNA sequence.
VERSION BQ923651
KEYWORDS BQ923651.1 GI:22338682
SOURCE EST.
ORGANISM human.
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
COMMENT NIH-MGC http://mgc.nci.nih.gov/
Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHM4009 row: e column: 02
High quality sequence stop: 583.
Location/Qualifiers
1. 901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6473545"
/lssue_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 266 a 181 c 234 g 220 t
ORIGIN
Query Match 20.3%; Score 659.6; DB 14; Length 901;
Best Local Similarity 87.0%; Pred. No. 9.8e-161;
Matches 840; Conservative 0; Mismatches 14; Indels 112; Gaps 6;
QY 155 TCTGGAAACTACAGTCTATGCTTTGAAGCGCAAAAGGGAATTAACATTTAAGACTCCC 214
Db 7 TCTGGAACACTACAGTCTATGCTTTGAAGCGCAAAAGGGAATTAACATTTAAGACTCCC 66
QY 215 CGGGAGCTGGAGATGAGCTTTCCATGCTGGCGGAGCGAGCAGCTTACATGAAAAT 274
Db 67 CGGGAGCTGGAGATGAGCTTTCCATGCTGGCGGAGCGAGCAGCTTACATGAAAAT 125
QY 275 CAGTAGATTAACCTCGCTCAGCTCTGTTTCAGAAAGTCTTTGCTCAGATCTTCTC 334
Db 126 ----- 125
QY 335 AGTGAGAAAGGGAAATACAGAAAGATGTCCTCCCTACAGTGTCTCAGAGACTGGTCTC 394
Db 126 -----TCAGACACTGGTCTC 141
QY 395 TTGGAGAAACTATAGTGGCAAAATCCCATTAACACCAATGACTTCAAAATTTTAAAAA 454
Db 142 TTGGAGAAACTATAGTGGCAAAATCCCATTAACCAATGACTTCAAAATTTTAAAAA 201
QY 455 ATAATGACCGTACGCTGTGTAAGCTCCCAAGTAATAGTTGGTGTATCTACCCCGG 514
Db 202 ATAATGACCGTACGCTGTGTAAGCTCCCAAGTAATAGTTGGTGTATCTACCCCGG 261
QY 515 TCTCTCAGTTTCAGGAAGCAAGCAAAATCTGCAAGTGTTCAGAAAAATGCTGACTC 574
Db 262 TCTCTCAGTTTCAGGAAGCAAGCAAAATCTGCAAGTGTTCAGAAAAATGCTGACTC 321
QY 575 CTAGAGTAGAGTTATCACTGTGGAAGATGACCTCACCACATGCTGTTATGCTGTGG 634
Db 322 CTAGAGTAGAGTTATCACTGTGGAAGATGACCTCACCACATGCTGTTATGCTGTGG 381

```

```

QY 635 TGAATGACCCCAATGAGATCTTCTGCATGGGGAGGCTGGCCCTGGCTTAAG 694
Db 382 TGATGACGCAATGAGATCTTCTGCATGGGGAGGCTGGCCCTGGCTTAAG 441
QY 695 CTGCTGATTTTGAATCCAGAGAGCAAAAGCTTTGTTGCCAGATATGTAAGTGT 754
Db 442 CTGCTGATTTTGAATCCAGAGAGCAAAAGCTTTGTTGCCAGATATGTAAGTGT 501
QY 755 CAGCTGTGATATGCTGTACAGGAGACAGGAGGCTCCCTGCAACAGATCATCATG 814
Db 502 CAGCTGTGATATGCTGTACAGGAGACAGGAGGCTCCCTGCAACAGATCATCATG 561
QY 815 CTGTTGGGCCCTGGGATGGAATGGATTAACAGGATGACTGGAAGCTGAGAGGG 874
Db 562 CTGTTGGGCCCTGGGATGGAATGGATTAACAGGATGACTGGAAGCTGAGAGGG 621
QY 875 CCATTTGATATTTCTGAATTTATGCTATTAATAAAATCTACATTAAGACAGTAA 934
Db 622 CCATTTGATATTTCTGAATTTATGCTATTAATAAAATCTACATTAAGACAGTAA 680
QY 935 TTCCAGCCTTGAAGCTTGGGATTTTTCAGTTCCTCTGATTTTGTACAAAGCTATTG 994
Db 681 TTCCAGCCTTGAAGCTTGGGATTTTTCAGTTCCTCTGATTTTGTACAAAGCTATTG 740
QY 995 TAGAGACTAT-CCGGGTAGTTTTCAGAGGAGCAATGATGA-GTAATTTGAAGAAT 1052
Db 741 TAGAGACTATCCGGGTAGTTTTCAGAGGAGCAATGATGATGATTTGAAGAAT 800
QY 1053 TCACCTGGTGAGCAAT-GAGGACCTTACTGTTG--CTGCTTTAAGCTGCTTACAGAT 1108
Db 801 TCACCTGGTGAGCAATGAGGACCTTACTGTTGCTGCTTTAAGCTGCTTCAAAAT 860
QY 1109 TCATCC 1114
Db 861 TCTCTC 866

```

RESULT 8
LOCUS BG284006 791 bp mRNA linear EST 21-FEB-2001
DEFINITION 602407302F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4519248 5',
mRNA sequence.
ACCESSION BG284006
VERSION BG284006.1 GI:13034520
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTD/DPF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHM40415 row: d column: 01
High quality sequence stop: 729.
Location/Qualifiers
1. 791
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4519248"
/lssue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

Db 548 AATATGTACAGTACACAGATTAATCAATCAGACCAATGACAGCCCTTGGACAGATCCCTT 489
QY 2759 GGAGGGATTTGCAAGTGGACAGCCCTGTTGATTAATCTCATATCATTTTAACAGCTGGT 2818
Db 488 GGAGGGATTTGCAAGTGGACAGCCCTGTTGATTAATCTCATATCATTTTAACAGCTGGT 429
QY 2819 ATGGCTTACCTGGGTGAGTAAACCAAAATATGACATGAGGCTCAAGAGTGGCTT 2878
Db 428 ATGGCTTACCTGGGTGAGTAAACCAAAATATGACATGAGGCTCAAGAGTGGCTT 369
QY 2879 GAATATATCCCATGGGTATATCTGATGACATGACGCGGTATGAAAGACTAGCCACAT 2938
Db 368 GAATATATCCCATGGGTATATCTGATGACATGACGCGGTATGAAAGACTAGCCACAT 309
QY 2939 ACTAGCATCTAGTACCTTATCTGATGACATGACGCGGTATGAAAGACTAGCCACAT 2998
Db 308 ACTAGCATCTAGTACCTTATCTGATGACATGACGCGGTATGAAAGACTAGCCACAT 249
QY 2999 AATGAAACACTTTGAGACAGCTTCCCTCTGAGAGTGTCTTAAATCTCTTACTAG 3058
Db 248 AATGAAACACTTTGAGACAGCTTCCCTCTGAGAGTGTCTTAAATCTCTTACTAG 189
QY 3059 AGGAGATTAATATTTTGCATATATGAAGAATTTTCTAGTATATTAACGACAGCCCTTT 3118
Db 188 AGGAGATTAATATTTTGCATATATGAAGAATTTTCTAGTATATTAACGACAGCCCTTT 129
QY 3119 ATTTCTTAATATGATGATAGATATAAATGTTAGATTAACGAATGATTTAGATTTCC 3178
Db 128 ATTTCTTAATATGATGATAGATATAAATGTTAGATTAACGAATGATTTAGATTTCC 69
QY 3179 AGAGATTAATTAAGTCTTAGTATGAAGAATTAATCATCTTGTCTGATTAATAAAA 3238
Db 68 AGAGATTAATTAAGTCTTAGTATGAAGAATTAATCATCTTGTCTGATTAATAAAA 9
QY 3239 AAAAA 3243
Db 8 AAAAA 4

RESULT 10
LOCUS B0028103/c 658 bp mRNA linear EST 27-MAR-2002
DEFINITION UI-H-CO0-ari-b-03-0-UI.s1 NCI_CGAP_Sub9 Homo sapiens cDNA clone
IMAGE:3106491.3, mRNA sequence.
ACCESSION B0028103
VERSION B0028103.1 GI:19763382
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eumetazoa; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 658)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL J. Biol. Chem. 272(19):11580-11586 (1997).
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
PolA-tails.

FEATURES
Source 1. .658 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3106491"
/clone_lib="NCI_CGAP_Sub9"
/tissue_type="mixed"

/dev_stage="mixed"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker. Site 1: EcoR I; Site 2: Not I; tissues:
Cholonic mucosa with Crohn's disease, Cholonic mucosa with
ulcerative colitis, Fetal thymus, Cervix, Cervical
adenosquamous carcinoma, Ligament cells, Prostate
carcinoma, Bladder carcinoma, Brain oligodendrocyte ;
NCI CGAP Sub9 is a subcloned cDNA library constructed
according to Bonaldo, Lemon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are GTC, AACG, GGGCC,
GGAG, TACG, TAAGC, ATGG, AGACA, ATGAC. For additional
information, contact: Bento Soares, bento-soares@uiowa.edu
TAG-LIB-UI-H-CO0
TAG-TISSUE=Prostate Carcinoma
TAG_SEQ=ATGG*

BASE COUNT 147 a 156 c 118 g 235 t 2 others
ORIGIN

Query Match 19.9% Score 645; DB 14; Length 658;
Best Local Similarity 98.9%; Pred. No. 5.6e-157;
Matches 648; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1639 ACAGATTTGGAGATATATTAAGGCTTTAGTCTGAATGCGAAGAGTCCCAAGATGCTG 1698
Db 658 ACAGATTTGGAGATATATTAAGGCTTTAGTCTGAATGCGAAGAGTCCCAAGATGCTG 599
QY 1699 AGTTGAACAATTAACAGTGTCCCAAGTCAACAGAGAGGAGAAAGAAATGGGCTT 1758
Db 598 AGTTGAACAATTAACAGTGTCCCAAGTCAACAGAGAGGAGAAAGAAATGGGCTT 539
QY 1759 GAAGTATATCTCTGCGCATATCTGATGATGATTAACAGTGAAGAGATATGAGGCC 1818
Db 538 GAAGTATATCTCTGCGCATATCTGATGATGATTAACAGTGAAGAGATATGAGGCC 479
QY 1819 CAGCATGATCAAGAAAGATCTGAGTCCAGAACCCACATCATGGAATATCAT 1878
Db 478 CAGCATGATCAAGAAAGATCTGAGTCCAGAACCCACATCATGGAATATCAT 419
QY 1879 ATTTGTACCTGGAGAGAAAGATGACATTTTGTCTGAGCTTCAGAAAATCTCAAGT 1938
Db 418 ATTTGTACCTGGAGAGAAAGATGACATTTTGTCTGAGCTTCAGAAAATCTCAAGT 359
QY 1939 GTCTCCATCAGAAATATTAAGCCAGAGAGGACAGATTAAGATTAAGAGGCCCGG 1998
Db 358 GTCTCCATCAGAAATATTAAGCCAGAGAGGACAGATTAAGATTAAGAGGCCCGG 299
QY 1999 GCTGACCTCATAGGAGGTGATTAAGCAATTAAGATTAAGATTAAGATTAAGATTAAG 2058
Db 298 GCTGACCTCATAGGAGGTGATTAAGCAATTAAGATTAAGATTAAGATTAAGATTAAG 239
QY 2059 ATGGCAAGAAAGAGAGCCCTTGGCGCTGCTTGAAGACAGTGAATTAAGCA 2118
Db 238 ATGGCAAGAAAGAGAGCCCTTGGCGCTGCTTGAAGACAGTGAATTAAGCA 179
QY 2119 CAAAAAACCCAAAGAGAAATGAAGAAATATTAAGTGAAGATGCTGCTGCTCA 2178
Db 178 CAAAAAACCCAAAGAGAAATGAAGAAATATTAAGTGAAGATGCTGCTGCTCA 119
QY 2179 ACTCAAGAGCTTATGATCAAAAGAAACAGTTTGAAGATGCTGCTGCTTAAG 2238
Db 118 ACTCAAGAGCTTATGATCAAAAGAAACAGTTTGAAGATGCTGCTGCTTAAG 59
QY 2239 GTGGAAGAGATGACATGAGGCTCTTATGGCTGCTTCAAGAGAGAGAGAA 2293
Db 58 GTGGAAGAGATGACATGAGGCTCTTATGGCTGCTTCAAGAGAGAGAGAGAA 4

RESULT 11
BM829571
LOCUS BM829571 608 bp mRNA linear EST 06-MAR-2002
DEFINITION K-EST0102596 S9SNU601 Homo sapiens cDNA clone S9SNU601-59-H09 5', mRNA sequence.
ACCESSION BM829571
VERSION BM829571.1 GI:19185980
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 608)
 Kim, N.S., Hahn, Y., Oh, J. H., Lee, J. Y., Ahn, H. Y., Chu, M. Y., Kim, M. R., Oh, K. J., Cheong, J. E., Sohn, H. Y., Kim, J. M., Park, H. S., Kim, S. and Kim, Y. S.
AUTHORS Kim, Y. S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 59 row: H column: 09
 High quality sequence stop: 608.
FEATURES
 source
 1..608
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="S9SNU601-59-H09"
 /clone_lib="S9SNU601"
 /sex="M"
 /tissue_type="ascites"
 /cell_type="Epithelial"
 /cell_line="SNU-601"
 /lab_host="Top10F"
 /note="Organ: Stomach; Vector: pME18-Fl3; Site.1: XhoI; Site.2: XhoI. The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME185-Fl3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 196 a 114 c 132 g 166 t
ORIGIN
 Query Match 18.6%; Score 604.8; DB 14; Length 608;
 Best Local Similarity 99.7%; Pred. No. 1.7e-146;
 Matches 606; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 981 TTCAAGACATATGTAGAGACTTCGGGGTTAGTTGCAAGGAGCAATGATGAGTAA 1040
 DB 1 TCAAAAGACTATTGTAGAGACTATCCGGGTAGTTGCAAGGAGCAATGATGAGTAA 60
 QY 1041 TTGGAAGAATTCACCTGGTGAAGCAATGAGACCCCTACTGTGTGCTTCTTAAAGCTGC 1100
 DB 61 TTGGAAGAATTCACCTGGTGAAGCAATGAGACCCCTACTGTGTGCTTCTTAAAGCTGC 120
 QY 1101 TTGACATTCATCTGAGGAGAGTGAAGTGGAGCAAGAACCCCTTCTTCAATGC 1160
 DB 121 TTCAGATTCATCTGAGGAGAGTGAAGTGGAGCAAGAACCCCTTCTTCAATGC 180

QY 1161 AATGTCGTGAACCAACCTGACCCCTCCAGATTGTCCAGGGCCACATTGAATGGCAGACGGC 1220
 DB 181 AATGTCGTGAACCAACCTGACCCCTCCAGATTGTCTGGCCACATTGAATGGCAGACGGC 240
 QY 1221 AGATGTAATTTGTAATTTCTGTGAATACCCACATGATATTACAGTTGGACCTGTGGCAAGTC 1280
 DB 241 AGATGTAATTTGTAATTTCTGTGAATACCCACATGATATTACAGTTGGACCTGTGGCAAGTC 300
 QY 1281 AATTCACAAACAGCAGGAGTGAATGAATGAATTCGTAATTCCTGGCACAAGGCTAAACA 1340
 DB 301 AATTCACAAACAGCAGGAGTGAATGAATGAATTCGTAATTCCTGGCACAAGGCTAAACA 360
 QY 1341 GTTTCACAGCTCCCACTGATGCTACTGCTCAAAAAGATTAACTTTGTAATATATAT 1400
 DB 361 GTTTCACAGCTCCCACTGATGCTACTGCTCAAAAAGATTAACTTTGTAATATATAT 420
 QY 1401 ATACCATGATCTGTGCTGATTCAGAAATTTCTTAACCTCGATATTAACATGCAATGAA 1460
 DB 421 ATACCATGATCTGTGCTGATTCAGAAATTTCTTAACCTCGATATTAACATGCAATGAA 480
 QY 1461 GGAGCTTTGGAAAATATGCAATGAGCAAAATATTAATTCATTTCTTCCTGCTGCTGG 1520
 DB 481 GGAGCTTTGGAAAATATGCAATGAGCAAAATATTAATTCATTTCTTCCTGCTGCTGG 540
 QY 1521 GACTGGAACATGGAATTAAGAGGAACAGCAGCAGAGATTGTTGATGAAGTTT 1580
 DB 541 GACTGGAACATGGAATTAAGAGGAACAGCAGCAGAGATTGTTGATGAAGTTT 600
 QY 1581 AACATTG 1588
 DB 601 AACATTG 608

RESULT 12
BI857151
LOCUS BI857151 705 bp mRNA linear EST 10-OCT-2001
DEFINITION 603383747F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5392715 5', mRNA sequence.
ACCESSION BI857151
VERSION BI857151.1 GI:15997898
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 705)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@rs-remail.nih.gov
 Tissue Procurement: DCTD/DMP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LNA12000 row: n column: 12
 High quality sequence stop: 691.
FEATURES
 source
 1..705
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5392715"
 /clone_lib="NIH_MGC_87"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: breast; Vector: pCMV-Sport6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies."

BASE COUNT 226 a 121 c 151 g 207 t
 ORIGIN

Note: this is a NIH-MGC Library."

Query Match 18.5%; Score 600.4; DB 13; Length 705;
 Best Local Similarity 98.4%; Pred. No. 2.5e-145;
 Matches 659; Conservative 0; Mismatches 6; Indels 5; Gaps 5;

QY 2576 CACGAGACCTGGAGCTATACATGCTGATGACAGTGTGGTGGACATGCTCCAGCCCTG 2635
 Db 1 CACGAGACCTGGAGCTATACATGCTGATGACAGTGTGGTGGACATGCTCCAGCCCTG 60
 QY 2636 AACCTTGTATTTTATGTCATGTCAGTCATACCTCAGTATTTGGATGACATGACCC 2695
 Db 61 AAAC- TTGTATTTTATGTCATGTCAGTCATACCTCAGTATTTGGATGACATGACCC 118
 QY 2696 AGGAATATGACAGTACACAGATTACTCATGAGCCAAATGAGACCTTTGCACAGATC 2755
 Db 119 AGGAATATGACAGTACACAGATTACTCATGAGCCAAATGAGACCTTTGCACAGATC 178
 QY 2756 CTGGAGAGGATTTGGCAAGTGGCAGCCCTGTGATTTATCTCTACATCTTTTAACACT 2815
 Db 179 CTGGAGAGGATTTGGCAAGTGGCAGCCCTGTGATTTATCTCTACATCTTTTAACACT 238
 QY 2816 GGTATGGCTTACCTTGGGTCAGTACCAATATATGACATGTCATGCTCAAGAGTGG 2875
 Db 239 GGTATGGCTTACCTTGGGTCAGTACCAATATATGACATGTCATGCTCAAGAGTGG 298
 QY 2876 CTGGAATATATCCATGGGTTATCTGTATGAGTACTGAGTGGTTATTTGAAGACTAGCCA 2935
 Db 299 CTGGAATATATCCATGGGTTATCTGTATGAGTACTGAGTGGTTATTTGAAGACTAGCCA 358
 QY 2936 CATCTACATCTTAGTCCCTTATCTCTTATGTCCTTATGTCCTTGGGTCAGTACGATA 2995
 Db 359 CATCTACATCTTAGTCCCTTATCTCTTATGTCCTTATGTCCTTGGGTCAGTACGATA 418
 QY 2996 CCAATATGAACACTTTCAGGACCTTCTCTTCTGACGTTGCTTAAATCTCCTTAC 3055
 Db 419 CCAATATGAACACTTTCAGGACCTTCTCTTCTGACGTTGCTTAAATCTCCTTAC 478
 QY 3056 TAGAGGAGATTAATTTTTCATATATGAGAAATTTTCTAGTATTAACGAGGCT 3115
 Db 479 TAGAGGAGATTAATTTTTCATATATGAGAAATTTTCTAGTATTAACGAGGCT 537
 QY 3116 TTTATTTCTTAAATGATGATAGTAT -AAAATGTTAGGATTAACGATGATTTTAGATT 3174
 Db 538 TTTATTTCTTAAATGATGATAGTATTAATAAATGTTAGGATTAACGATGATTTTAGATT 597
 QY 3175 TTCCAGAAATATTAATAAGTCTTTAGTATGAAAT -AAATCATCTTTGCTGATTAA 3233
 Db 598 TTCCAGAAATATTAATAAGTCTTTAGTATGAAATTAATAATCATCTTTGCTGATAAAA 657
 QY 3234 AAAAAAAAAA 3243
 Db 658 AAAAAAAAAA 667

RESULT 13
 AM851285/c 601 bp mRNA linear EST 19-MAY-2000
 LOCUS AM851285
 DEFINITION IL3-CT0220-170200-067-H02 CT0220 Homo sapiens CDNA, mRNA sequence.
 ACCESSION AM851285
 VERSION AM851285.1 GI:7946802
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 601)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

TITLE 'M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 JOURNAL Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 MEDLINE Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 COMMENT 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?IL3-CT0220-170
 200-067-H02&ts=2000-02-17&tl=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 600.

FEATURES
 source
 Location/Qualifiers
 1..601
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0220"
 /dev_stage="adult"
 /note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:
 Sma1; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 ,716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."

BASE COUNT 191 a 131 c 123 g 156 t
 ORIGIN

Query Match 18.1%; Score 586.8; DB 10; Length 601;
 Best Local Similarity 99.5%; Pred. No. 8.4e-142;
 Matches 599; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2506 TATGTGTTGAGGCTGAAGTACTACAGGCTTCTTCCAGGAGCATCCGTTAAATATT 2565
 Db 601 TATGTGTTGAGGCTGAAGTACTACAGGCTTCTTCCAGGAGCATCCGTTAAATATT 542
 QY 2566 GTTCCCCACACATGACGCTCGAGGCTATACATGCTATGACAGTGTGTTGAACAATGTC 2625
 Db 541 GTT-CCCCACACATGACGCTCGAGGCTATACATGCTATGACAGTGTGTTGAACAATGTC 483
 QY 2626 TCCAGCCCTGAACCTTGTATTTTATGTCGATGACGAGCTATACCTCAGTATTTGTGG 2685
 Db 482 TCCAGCCCTGAACCTTGTATTTTATGTCGATGACGAGCTATACCTCAGTATTTGTGG 423
 QY 2686 ACATGACCCACAGAAATATGACAGTACACAGATTATCATCAGGACCAATGAGACCTTT 2745
 Db 422 ACATGACCCACAGAAATATGACAGTACACAGATTATCATCAGGACCAATGAGACCTTT 363
 QY 2746 GCACAGCATCCCTGGAGGGGAGTTGCGAAGTGGAGCCCTGTTGANTATCTCTCATCAT 2805
 Db 362 GCACAGCATCCCTGGAGGGGAGTTGCGAAGTGGAGCCCTGTTGANTATCTCTCATCAT 303
 QY 2806 TTTAAGACCTGTATGAGCTTACCTTGGGTGAACATAACCAATATATGACATGATGCT 2865
 Db 302 TTTAAGACCTGTATGAGCTTACCTTGGGTGAACATAACCAATATATGACATGATGCT 243
 QY 2866 CAAGAGTGGCTTGAATATATCCATGAGGTTATCTGTATGAGATGACTGGTTATTGAAA 2925
 Db 242 CAAGAGTGGCTTGAATATATCCATGAGGTTATCTGTATGAGATGACTGGTTATTGAAA 183
 QY 2926 GGACTAGCACAATATACATCTTAGTCCCTTATCTCTTTTATGTCCTTTGGGGTGGGG 2985
 Db 182 GGACTAGCACAATATACATCTTAGTCCCTTATCTCTTTTATGTCCTTTGGGGTGGGG 123
 QY 2986 TAGGTAGTACCACCAATGAAACACTTCAGACCTTCTTCTTGCAGTGTTCCTTTAA 3045

```
|||||
Db 122 TAGAGTACACCAATGAACATTTTCAGACCTTCCTCTTCAGATGTCTTTATA 63
OY 3046 TCTCCCTTACTAGAGAGATTAATTTTGCATATATATGAGAAATTTTCTAGTATA 3105
Db 62 TCTCCCTTACTAGAGAGATTAATTTTGCATATATATGAGAAATTTTCTAGTATA 3
OY 3106 AC 3107
Db 2 AC 1
RESULT 14
LOCUS B1836699 732 bp mRNA linear EST 04-OCT-2001
DEFINITION 603089634F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5228481 5',
mRNA sequence.
ACCESSION B1836699
VERSION B1836699.1 GI:15948249
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 732)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11575 row: C column: 10
High quality sequence stop: 732.
Location/Qualifiers
1. 732
source /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5228481"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
PCMV-SPORT6; Site_1: NOTI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 023. Note: this is a NIH-MGC Library."
BASE COUNT 213 a 150 c 187 g 182 t
ORIGIN
Query Match 18.1%; Score 586.2; DB 13; Length 732;
Best Local Similarity 86.7%; Pred. No. 1.3e-141;
Matches 726; Conservative 0; Mismatches 3; Indels 106; Gaps 3;
OY 137 CAGAGTGGGCGCCATATCTGGAACACTACAGTCTATGCTTTGAAGCGCAAAAGGAAAT 196
Db 1 CGGAAGTGGCGCCACATATCTGGAACACTACAGTCTATGCTTTGAAGCGCAAAAGGAAAT 60
OY 197 AACATTTAAAGACTCCCCCGGGAGCTGGAGATGAGCTTTTCATGTGTGGCGGAGAC 256
Db 61 AACATTTAAAGACTCCCCCGGGAGCTGGAGATGAGCTTTTCATGTGTGGCGGAGAC 120
OY 257 CAGCTTACATTAATAATAGGTAGATTAAGTCTGCTCACTCTTGTTCAGAAAGTCT 316
Db 121 CAGCTTACATTAATAATAA----- 137
```

```
OY 317 TTGCTCAGATCTTCTCCTCAGTGGAGAAAGGGAATACAGAAATGTCTCCCTACAGT 376
Db 138 ----- 137
OY 377 GCTCAGAGACTGTGCTCTTGGAGAAACTATAGTTGGCAATTCOCATTAAACCAATG 436
Db 138 --TCAGAGACTGTGCTCTTGGAGAAACTATAGTTGGCAATTCOCATTAAACCAATG 195
OY 437 ACTTCAA-AATTTTAAATAATAGAGCGTCAGCTGTGTGAAGTCCCTCAGATAAGTTT 495
Db 196 ACTTCAACAATTTTAAATAATAGAGCGTCAGCTGTGTGAAGTCCCTCAGATAAGTTT 255
OY 496 GCGTGTATCTTACCTGTGTCTCTCCAGTTCAAGGAAGCAACAGCAATCTGCAAGTG 555
Db 256 GCGTGTATCTTACCTGTGTCTCTCCAGTTCAAGGAAGCAACAGCAATCTGCAAGTG 315
OY 556 TTCCA--GAAAAATGCTGATCCTAGATAGATTTATCACTGCGAAAGTGAACCTCACA 613
Db 316 TTCAATGAATAATTAATGCTGACTCTAGATAGATTTATCACTGCGAAAGTGAACCTCACA 375
OY 614 CACATGCTGTGATGCTGTGTGAATGCAGCAATGAATCTTCTGCATGGGGAGGCC 673
Db 376 CACATGCTGTGATGCTGTGTGAATGCAGCAATGAATCTTCTGCATGGGGAGGCC 435
OY 674 TGGCCCTGGCCCTGTGTAAGAGCTGTGATTTGAATCCAGAGAGACCAAGTTTG 733
Db 436 TGGCCCTGGCCCTGTGTAAGAGCTGTGATTTGAATCCAGAGAGACCAAGTTTG 495
OY 734 TTGCCAGATATGTGTAAGAGCTGTGATGATAGCTGTGATGATGATGATGATGATGAT 793
Db 496 TTGCCAGATATGTGTAAGAGCTGTGATGATGATGATGATGATGATGATGATGATGAT 555
OY 794 CCTGCAACAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 853
Db 556 CCTGCAACAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 615
OY 854 GTACTGGAAGCTGCAGAGAGGCAATTTGATGATGATGATGATGATGATGATGATGATGAT 913
Db 616 GTACTGGAAGCTGCAGAGAGGCAATTTGATGATGATGATGATGATGATGATGATGATGAT 675
OY 914 CTCACATTAAGACAGTACGATATTCACAGCTCTGAGCTCTGAGTTTTCAGTTCCTCC 970
Db 676 CTCACATTAAGACAGTACGATATTCACAGCTCTGAGCTCTGAGTTTTCAGTTCCTCC 732
RESULT 15
LOCUS AM851000 610 bp mRNA linear EST 19-MAY-2000
DEFINITION IL3-CT0220-150200-069-E07 CT0220 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM851000
VERSION AM851000.1 GI:7946517
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brites,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,L.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
```

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=6t2-IL3-CT0220-150
200-069-E07&L3=2000-02-15&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 545.

FEATURES

Source

1. 610
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CF0220"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
, 716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 208 a 114 c 147 g 140 t 1 others
ORIGIN

Query Match 18.0%; Score 584.2; DB 10; Length 610;
Best Local Similarity 98.2%; Pred. No. 4e-141;

Matches 601; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

```
OY 1776 CATCAATCTGATGGATTCACAGTGAAGAGATGTATGAGGCCACGCGATGGATCCAAAG 1835
|||||
DB 1 CATCAATCTAATGGATTCAAGCTGAAGAGATGTATGAGGCCACGCGATGGATCCAAAG 60
OY 1836 AATCTGAGTCTCCGAACCCACACATCATTTGAGAAATATCATATCTGTACCTTGGCAG 1895
|||||
DB 61 AATCTGAGTCTCCGAACCCACACATCATTTGAGAAATATCATATCTGTACCTTGGCAG 120
OY 1896 AAAGGACATGATCTTTGTTCTCAGCTTCAGAAACTTCAGAGTCTCCATCAGCAAGAAAT 1955
|||||
DB 121 AAAGGACATGATCTTTGTTCTCAGCTTCAGAAACTTCAGAGTCTCCATCAGCAAGAAAT 180
OY 1956 TATCAGCCAGGAGAGACAGATTGAGATTGAAAGAGCCCGGCTGACCTCATTTGAGGT 2015
|||||
DB 181 TATCAGCCAGGAGAGACAGATTGAGATTGAAAGAGCCCGGCTGACCTCATTTGAGGT 240
OY 2016 GGTATGACATTTGAAGATATGCTTTGTAAGTACAGAGGAAATGCAAGGAAAAAGGA 2075
|||||
DB 241 GGTATGACATTTGAAGATATGCTTTGTAAGTACAGAGGAAATGCAAGGAAAAAGGA 300
OY 2076 GCGAGGCGTTGGGCTGTTGGACAGTGGACTATTTCAGCAACAAAAAACCCAGAGCA 2135
|||||
DB 301 GCGAGGCGTTGGGCTGTTGGACAGTGGACTATTTCAGCAACAAAAAACCCAGAGCA 360
OY 2136 AATGAAGAAAAATATCATATTTCTGAATATGCTGCTGCCACTCAAGAGCTTCTAGA 2195
|||||
DB 361 AATGAAGAAAAATATCATATTTCTGAATATGCTGCTGCCACTCAAGAGCTTCTAGA 420
OY 2196 TCAAAAGAAACAGTTGAAAAATGTGTTGAGGTTCTAAAGTGGAGAGATAGACAA 2255
|||||
DB 421 TCAAAAGAAACAGTTGAAAAATGTGTTGAGGTTCTAAAGTGGAGAGATAGACAA 480
OY 2256 TGAGGTCTTATGGCTGCTTCAAGAAAGAAAGAAATGATGSAAGAAAACTGCACAG 2315
|||||
DB 481 TGAGGTCTTATGGCTGCTTCAAGAAAGAAAGAAATGATGSAAGAAAACTGCACAG 538
OY 2316 GCACCTGTGAGCCATAGGCTGTTCAGCAAGTCCCATACAGTCTGCAATGTGTATG 2375
|||||
DB 539 GCACCTGTGAGCCATAGGCTGTTCAGCAAGTCCCATACAGTCTGCAATGTGTATG 598
OY 2376 CAGAGTTGGCTT 2387
|||||
DB 599 CAGAGTTGGCTT 610
```

Search completed: February 7, 2003, 05:35:47
Job time : 2907 secs

THIS PAGE BLANK (USPTO)

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

17257.341 Million cell updates/sec

Sequence: 1 gggtctcgtgttcctggtg.....gtctgatttaaaaaaaaaa 3243

Scoring table: IDENTITY_NUC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database

1:	gb_ba:*
2:	gb_hng:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vl:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_jmu:*
19:	em_om:*
20:	em_or:*
21:	em_ov:*
22:	em_pat:*
23:	em_ph:*
24:	em_pl:*
25:	em_ro:*
26:	em_sts:*
27:	em_un:*
28:	em_vl:*
29:	em_vl:*
30:	em_hng:*
31:	em_hng_in:*
32:	em_hng_other:*
33:	em_hng_mus:*
34:	em_hng_pla:*
35:	em_hng_rod:*
36:	em_hng_mam:*
37:	em_hng_vre:*
38:	em_sy:*
39:	em_hng_hum:*
40:	em_hng_mus:*
41:	em_hng_ob:*

Pred. No. is the number of results predicted by chance to have a

Result	Score	Query Match	Length	DB	ID	Description
No. <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>						
1	3241.4	100.0	3243	9	AF307338	AF307338 Homo sapi
2	3021.4	93.2	3138	9	AF307339	AF307339 Homo sapi
3	2842.6	78.7	3016	6	AX399212	AX399212 Sequence
4	2543.8	87.4	2598	6	AX274946	AX274946 Sequence
5	1440.4	44.4	1442	9	HSM803023	AL713679 Homo sapi
6	1371.8	42.3	1388	9	HSM804240	AL832929 Homo sapi
7	1362.4	42.0	3012	10	BC003281	BC003281 Mus muscu
8	844.4	26.0	167164	2	AC026506	AC026506 Homo sapi
9	844.4	26.0	174257	2	AC092908	AC092908 Homo sapi
C 10	821	25.3	167164	2	AC026506	AC026506 Homo sapi
C 11	632.8	19.2	822	6	AX188088	AX188088 Sequence
C 12	489.4	15.1	666	6	AX188241	AX188241 Sequence
C 13	463.2	14.3	516	6	AX185819	AX185819 Sequence
C 14	449.8	13.9	481	6	AX187250	AX187250 Sequence
C 15	414.8	12.8	457	6	AX184391	AX184391 Sequence
C 16	299	9.2	299	6	AX396731	AX396731 Sequence
C 17	284	8.8	524	6	AX332099	AX332099 Sequence
C 18	280.4	8.6	524	6	AX186264	AX186264 Sequence
C 19	246.8	7.6	260	6	AX330378	AX330378 Sequence
C 20	246.8	7.6	260	6	AX333102	AX333102 Sequence
C 21	236.6	7.3	124558	2	AC095176	AC095176 Rattus no
C 22	209.8	6.5	387	6	AX187472	AX187472 Sequence
C 23	202.8	6.3	214	6	AX302704	AX302704 Sequence
C 24	181.2	5.6	183	6	AX341343	AX341343 Sequence
C 25	140.4	4.3	5718	9	HSM804911	AL833598 Homo sapi
C 26	140	4.3	2223	9	AF484416	AF484416 Homo sapi
C 27	127.8	3.9	261	6	AX184601	AX184601 Sequence
C 28	122.8	3.8	2589	9	AK026003	AK026003 Homo sapi
C 29	122.8	3.8	2713	9	AK098816	AK098816 Homo sapi
C 30	122.8	3.8	5231	9	AB033094	AB033094 Homo sapi
C 31	122.8	3.8	149307	2	AC010858	AC010858 Homo sapi
C 32	122.8	3.8	203092	2	AC048348	AC048348 Homo sapi
C 33	122.8	3.8	319404	2	AC112130	AC112130 Homo sapi
C 34	120.6	3.7	66539	2	AC095519	AC095519 Rattus no
C 35	117.6	3.6	124585	2	AC0955176	AC0955176 Rattus no
C 36	82.6	2.5	10265	1	AE001727	AE001727 Thermotogu
C 37	79	2.4	277050	1	AB000989	AB000989 Sulfolobu
C 38	75.6	2.3	11109	1	AE009804	AE009804 Pyrobacul
C 39	68.4	2.1	66539	2	AC095519	Rattus no
C 40	68.2	2.1	96208	1	AL596174	AL596174 Listeria
C 41	68.2	2.1	311208	6	AX417049	AX417049 Sequence
C 42	68.2	2.1	349980	6	AX417036	AX417036 Sequence
C 43	65.8	2.0	3194	6	AX416716	AX416716 Sequence
C 44	65.4	2.0	7218	6	166494	166494 Sequence 14
C 45	62.6	1.9	223528	1	AL591984	AL591984 Listeria

FEATURES	source
REFERENCE	1. 3243
AUTHORS	Aguiar,R.C.T., yakushijin.Y. and Shipp,M.A.
TITLE	Direct Submission
JOURNAL	Submitted (21-SEP-2000) Adult Oncology, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
gene	Location/Qualifiers
CDS	1. 3243
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="3"
	/map="3q13-q21"
	1. 3243
	/gene="BAL"
	229. 2793
	/gene="BAL"
	/codon_start=1
	/product="B aggressive lymphoma long isoform"
	/protein_id="AAK07558.1"
	/db_xref="GI:12751139"
	/translation="MDSWVAGAAAYNEKSGRTISLSLFOKVRQIIPOMRKTEBCLPCKSETGALGENYSWQIPINHDLEILKNVNEQALEDLONKGCISTLVSPQEEKSLQVREKMLTPRIELSWPKDITLTAVDVAANAEIDLGLGALALVAGFGLIOESKOPVARYKGVAGEIAVTGAGRPCKQIIVAGVPMRPMKOCSTGLQALAI SLUNVYIKNTHIKIVAIIPALSPSGFQFPLNCTIYETIRVSLDQKPMNSLKEIHLVSNEDPTVAAPKASERITLKSLELGGQTTTSPFNAMVNNLTQVQSHIEQTDVIVNSVPHDITGPVAKSLILOQAGVEMKSEFLATKQFQORSOLVLTGFINLEKTYIHLWMSERPQKQILKHAMKECLEKEQNTITSISFALPGQIMETIKETAEILFDEYLTFAKDHVQHLTVKVFYFEPDLEIKAFSSSMARKSLNNYVSQSTEEKREHGLEKSPAINLMGEVEMCEAHAIQRLISLQNHIIENNHILIGRKHILSOLKTSYSITSEISPGTELEIGARADLILEYNNIDMLCYOEEMARKERGLMSLQWTLDOQKTODEMKNITLFCQAPPPQELLDDKQFQKGLQVYKVDINEVMAAFORKKRMEKRLHROPVSHRLFOQVPYPCNVYCVGQRMSTSPCDRYAGCITFTYLNKLNIAEAKKIIISADKLIVYFAEVLVIGFQCGHPLIVPPSLPGALIDGDSVVDVSSPTPEVIFEGMAIIPQYLTWCTDEYVQSDYSFGPMRPFQHPWRGFASSGSPVD"
BASE COUNT	989 a 647 c 756 g 851 t
ORIGIN	
Query Match	100.0%; Score 3241.4; DB 9; Length 3243;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 3242; Conservative	0; Mismatches 1; Indels 0; Gaps 0.
1	GGGCTTCGTCGTCGCGGTCGACCGTCACATCCCGCGCGCGGAGCACTTAGAGCTCT 60
1	GGGCTTCGTCGTCGCGGTCGACCGTCACATCCCGCGCGCGGAGCACTTAGAGCTCT 60
61	GGAAGTAGCTCCACAGCTCTTCGTCACGCGGAGGCGGAGCTTGTACACCGCAGCAGGA 120
61	GGAAGTAGCTCCACAGCTCTTCGTCACGCGGAGGCGGAGCTTGTACACCGCAGCAGGA 120
61	GGAAGTAGCTCCACAGCTCTTCGTCACGCGGAGGCGGAGCTTGTACACCGCAGCAGGA 120
121	GGGGGACGCGCGGCGCAGACAGTGGGCCACATATCTGCAAACTACAGTATGCTTTGA 180
121	GGGGGACGCGCGGCGCAGACAGTGGGCCACATATCTGCAAACTACAGTATGCTTTGA 180
121	GGGGGACGCGCGGCGCAGACAGTGGGCCACATATCTGCAAACTACAGTATGCTTTGA 180
181	ACGCGAAAAGGGAATTAACATTTAAAGACTCCCGCGGAGCCTGGAGAGATGACATTTCG 240
181	ACGCGAAAAGGGAATTAACATTTAAAGACTCCCGCGGAGCCTGGAGAGATGACATTTCG 240
241	ATGGTGGCGCGGAGCAGACACTTACATGAATAATCAGATGAGATTACGTCGGCTCAGTC 300
241	ATGGTGGCGCGGAGCAGACACTTACATGAATAATCAGATGAGATTACGTCGGCTCAGTC 300
301	TTGTTTCACAAAGTCTTTCTCAGATCTTTCTCCTCAGTGGAGAAAGGGAATACAGAAGAA 360
301	TTGTTTCACAAAGTCTTTCTCAGATCTTTCTCCTCAGTGGAGAAAGGGAATACAGAAGAA 360
361	TGTCCTCCCTACAGTGTCTCAGAGACTGGTGTCTTTGGAGAAAACATATAGTTGGCAATT 420
361	TGTCCTCCCTACAGTGTCTCAGAGACTGGTGTCTTTGGAGAAAACATATAGTTGGCAATT 420

QY	421	CCCATTTAACCAATGACTTCAAAATTTTAAAAATATAGCGTCAGCTGTGTGAAGTC	480
Db	421	CCCATTTAACCAATGACTTCAAAATTTTAAAAATATAGAGCGTCAGCTGTGTGAAGTC	480
QY	481	CTCCAGATAGTTGGCTGTATCTCTACCCTGTCTCCAGCTCAGGAAGGCAACGC	540
Db	481	CTCCAGATATAGTTGGCTGTATCTCTACCCTGTCTCCAGCTCAGGAAGGCAACGC	540
QY	541	AAATCTCTGCAAGGTTCACAAAAATGCTGACTCCTAGAGATATATCACTGTGANA	600
Db	541	AAATCTCTGCAAGGTTCACAAAAATGCTGACTCCTAGAGATATATCACTGTGANA	600
QY	601	GATGACCTCACACACATGCTGTGATGCTGTGTGATGCAATGACCCATGAAATCTCTG	660
Db	601	GATGACCTCACACACATGCTGTGATGCTGTGTGATGCAATGACCCATGAAATCTCTG	660
QY	661	CATGGGGAGGCGCTGGCCCTGGCCCTGGTAAAAAGCTGTGTGATTTGAATCCAGAAAG	720
Db	661	CATGGGGAGGCGCGCCCTGGCCCTGGTAAAAAGCTGTGTGATTTGAATCCAGAAAG	720
QY	721	AGCAAAAGATTGTTGGCAGATATATGTAAGTGTACAGTGGAGATATAGCTCACGGA	780
Db	721	AGCAAAAGATTGTTGGCAGATATATGTAAGTGTACAGTGGAGATATAGCTCACGGA	780
QY	781	GCAGGAGGCGCTTCCCTGCAAAACAGATCATCATGCTGTTGGCGCTCGGTGATGGAATG	840
Db	781	GCAGGAGGCGCTTCCCTGCAAAACAGATCATCATGCTGTTGGCGCTCGGTGATGGAATG	840
QY	841	GATTAACAGGAGATCTACTGGAAGGCTGCAAGGGCCATTGTAAATTTCTGAATTAATGTC	900
Db	841	GATTAACAGGAGATCTACTGGAAGGCTGCAAGGGCCATTGTAAATTTCTGAATTAATGTC	900
QY	901	ATCTATATAAAATACTACATTAGACAGATGACATTCACGCTTATAGCTCTGGAATTTT	960
Db	901	ATCTATATAAAATACTACATTAGACAGATGACATTCACGCTTATAGCTCTGGAATTTT	960
QY	961	CAGTTCCTCTGAATTTGTGTACAAAGACTATTGTAGAGACTATCCGGGTTAGTTTGCA	1020
Db	961	CAGTTCCTCTGAATTTGTGTACAAAGACTATTGTAGAGACTATCCGGGTTAGTTTGCA	1020
QY	1021	GGGAAGCCATGATGATTAATTTGAAAGAAATTCACCTGTGAGCAATGAGACCTTACT	1080
Db	1021	GGGAAGCCATGATGATTAATTTGAAAGAAATTCACCTGTGAGCAATGAGACCTTACT	1080
QY	1081	GTTGCTGCGCTTAAAGCTGCTTCAAGATTCATCTTAGGAAGAGTGAAGTGGGACABAA	1140
Db	1081	GTTGCTGCGCTTAAAGCTGCTTCAAGATTCATCTTAGGAAGAGTGAAGTGGGACABAA	1140
QY	1141	ACCAACCCCTCTTTCATGTCGAATGCTGTGTAACACCTGACCCCTCAGATTGTCCAGGGC	1200
Db	1141	ACCAACCCCTCTTTCATGTCGAATGCTGTGTAACACCTGACCCCTCAGATTGTCCAGGGC	1200
QY	1201	CACATTGATGAGCAGACGCGCAGATGTATTTGTTAATTCGTATAACCCACATGATATTACA	1260
Db	1201	CACATTGATGAGCAGACGCGCAGATGTATTTGTTAATTCGTATAACCCACATGATATTACA	1260
QY	1261	GTTGAGACCTGTGGCAAGTCAATTTTCAACAACAGCAGGAGTTGAATGAAATCGGAATTT	1320
Db	1261	GTTGAGACCTGTGGCAAGTCAATTTTCAACAACAGCAGGAGTTGAATGAAATCGGAATTT	1320
QY	1321	CTTGCCACAAAGGCTTAACAGATTTCACAAAGGCTCCCATTTGTACTGTGTACAAAGGATTT	1380
Db	1321	CTTGCCACAAAGGCTTAACAGATTTCACAAAGGCTCCCATTTGTACTGTGTACAAAGGATTT	1380
QY	1381	AACTTGTCTGTAAATATATATACCAATGATGCTGTGATTCAGAAATTTCTTAAACCTCAG	1440
Db	1381	AACTTGTCTGTAAATATATATACCAATGATGCTGTGATTCAGAAATTTCTTAAACCTCAG	1440
QY	1441	ATATTAAACATGCAATGAAGAGAGTGTGGGAAAAATGATTTGACCAAAATTTAACTTCC	1500
Db	1441	ATATTAAACATGCAATGAAGAGAGTGTGGGAAAAATGATTTGACCAAAATTTAACTTCC	1500

1501 ATTTCCTTCCCTGCGCTGGAGCTGGAACATGAAATTAAGAGAGAAACAGCAGCAGAG 1560
 1501 ATTTCCTTCCCTGCGCTGGAGCTGGAACATGAAATTAAGAGAGAAACAGCAGCAGAG 1560
 1561 ATTTTGTGATGAAGTTTAAACATTTGGCAAGACCATGTAAGAACCCAGTTAACTGTA 1620
 1561 ATTTTGTGATGAAGTTTAAACATTTGGCAAGACCATGTAAGAACCCAGTTAACTGTA 1620
 1621 AAATTTGATGATCTTCCAAACAGATTGGAGATATATTAAGGCTTTGAGTTCTTAATGGGA 1680
 1621 AAATTTGATGATCTTCCAAACAGATTGGAGATATATTAAGGCTTTGAGTTCTTAATGGGA 1680
 1681 AAGAGGTCACAGATCTGAGTTTGAACAATTACAGTGTCCCCAGTCAACAGAGAGAG 1740
 1681 AAGAGGTCACAGATCTGAGTTTGAACAATTACAGTGTCCCCAGTCAACAGAGAGAG 1740
 1741 AAAAGAGAAATGGCTTGAAAGCTAGATCTCTGCTCCATCATCTGATGGATTTCAACGTG 1800
 1741 AAAAGAGAAATGGCTTGAAAGCTAGATCTCTGCTCCATCATCTGATGGATTTCAACGTG 1800
 1801 GAAGAGATGATGAGGCGCCAGCATGATGCAAGATCCAGTCTCCAGAACCCAGCAG 1860
 1801 GAAGAGATGATGAGGCGCCAGCATGATGCAAGATCCAGTCTCCAGAACCCAGCAG 1860
 1861 ATCATTGAGAAATATCATATTTCTGACCTTGGAGAGAAAGAACATGACATTTTGTCTCAG 1920
 1861 ATCATTGAGAAATATCATATTTCTGACCTTGGAGAGAAAGAACATGACATTTTGTCTCAG 1920
 1921 CTTCAGAAAATCTCAAGTGTCTCCATCACAGAAATTTATCAGCCCGAGAGAGAGATTA 1980
 1921 CTTCAGAAAATCTCAAGTGTCTCCATCACAGAAATTTATCAGCCCGAGAGAGAGATTA 1980
 1981 GAGATTTGAAGAGGCGCGGCTGACCTCATTTGAGGTTATGATTAACATTAAGATATGCTT 2040
 1981 GAGATTTGAAGAGGCGCGGCTGACCTCATTTGAGGTTATGATTAACATTAAGATATGCTT 2040
 2041 TGTAAAGTACAGAGAGAAATGGCAAGAAAGAGAGCGAGCTTTGGCGCTCGTTAGGA 2100
 2041 TGTAAAGTACAGAGAGAAATGGCAAGAAAGAGAGCGAGCTTTGGCGCTCGTTAGGA 2100
 2101 CAGTGGACTATTCAGACACAAAAAACCCAGACGAATGAAAGAAATATCATATTTCTG 2160
 2101 CAGTGGACTATTCAGACACAAAAAACCCAGACGAATGAAAGAAATATCATATTTCTG 2160
 2161 AATGTCCTGTCCTCCAACTCAAGAGCTTATGATCAAAAGAAACAGTTGAAAAATGT 2220
 2161 AATGTCCTGTCCTCCAACTCAAGAGCTTATGATCAAAAGAAACAGTTGAAAAATGT 2220
 2221 GGTTCGAGGTTCTAAGGTGAGAGAAATGAGCAATGAGGCTTATGAGCTGCTTCA 2280
 2221 GGTTCGAGGTTCTAAGGTGAGAGAAATGAGCAATGAGGCTTATGAGCTGCTTCA 2280
 2281 AGAAGAGAGAAATATGATGAGAGAAACTGCAAGAGCAACCTGTGAGCCATAGGCTGTT 2340
 2281 AGAAGAGAGAAATATGATGAGAGAAACTGCAAGAGCAACCTGTGAGCCATAGGCTGTT 2340
 2341 CAGCAAGTCCCTACCACTTCCGCAATGTGATGAGAGGTTGGCTTTAAAGAAATGTAC 2400
 2341 CAGCAAGTCCCTACCACTTCCGCAATGTGATGAGAGGTTGGCTTTAAAGAAATGTAC 2400
 2401 TCGACACCTTGGATCCAAATATAGAGAGCTGATATCTTACCAAGAAACCTCAAAAAC 2460
 2401 TCGACACCTTGGATCCAAATATAGAGAGCTGATATCTTACCAAGAAACCTCAAAAAC 2460
 2461 CTGGCAGAGAGAGCCAGAGAAATCTCTGCTGCAAGATAGCTGATGTGTTGAGGCT 2520
 2461 CTGGCAGAGAGAGCCAGAGAAATCTCTGCTGCAAGATAGCTGATGTGTTGAGGCT 2520
 2521 GAAGTACTACAGAGCTTCTTTCGCAAGAGACATCCGTTAAATATGTGTTCCCCACACATG 2580
 2521 GAAGTACTACAGAGCTTCTTTCGCAAGAGACATCCGTTAAATATGTGTTCCCCACACATG 2580
 2581 AGTCTGAGAGCTATAGATGCTATGACAGTGTGTTGACAAATGTCTCCAGCCCTGAACCC 2640

2581 AGTCTGAGAGCTATAGATGCTATGACAGTGTGTTGACAAATGTCTCCAGCCCTGAACCC 2640
 2641 TTTGTTATTTTATAGTGCATGAGGCTATACCTCAATTTTGTGACATGACCCAGGAA 2700
 2641 TTTGTTATTTTATAGTGCATGAGGCTATACCTCAATTTTGTGACATGACCCAGGAA 2700
 2701 TATGTACAGTCAAGATTTACTCATCAGCAAGCAATGAGACCTTTGACAGCATCTTGG 2760
 2701 TATGTACAGTCAAGATTTACTCATCAGCAAGCAATGAGACCTTTGACAGCATCTTGG 2760
 2761 AGGGATTCGCAAGTGGAGGCGCCCTGTTGATATCTTCAATCATATTTTAAACAGCTGAT 2820
 2761 AGGGATTCGCAAGTGGAGGCGCCCTGTTGATATCTTCAATCATATTTTAAACAGCTGAT 2820
 2821 GGCCTTACCTTGGTGAAGTCAACCAATATATACATGATGCTGCTCAAGAGTGGCTTGA 2880
 2821 GGCCTTACCTTGGTGAAGTCAACCAATATATACATGATGCTGCTCAAGAGTGGCTTGA 2880
 2881 ATATATCCCATGGGTTATCTGATGAGTACTGAGTGGGTTATGAAGAGTACCATATAC 2940
 2881 ATATATCCCATGGGTTATCTGATGAGTACTGAGTGGGTTATGAAGAGTACCATATAC 2940
 2941 TAGCATCTTATGAGCTTATCTGATGAGTACTGAGTGGGTTATGAAGAGTACCATATAC 3000
 2941 TAGCATCTTATGAGCTTATCTGATGAGTACTGAGTGGGTTATGAAGAGTACCATATAC 3000
 3001 TGAACACCTTACAGACCTCTCTCTGTCAGATTTCTTAACTCTTACTAGAG 3060
 3001 TGAACACCTTACAGACCTCTCTCTGTCAGATTTCTTAACTCTTACTAGAG 3060
 3061 GAGATTAATATTTTGCATTAATGAAGAAATTTCTTATATATTAAGCAGGCTTTTAT 3120
 3061 GAGATTAATATTTTGCATTAATGAAGAAATTTCTTATATATTAAGCAGGCTTTTAT 3120
 3121 TTTCTAAATATGATAGTATGATTAATGATTTAGATTAACGAATGATTTAGATTTCCAG 3180
 3121 TTTCTAAATATGATAGTATGATTAATGATTTAGATTAACGAATGATTTAGATTTCCAG 3180
 3181 AGAATATTTTAAAGTCTTATGATTAATGATTTAGTGTCTGATTTAAAAAAA 3240
 3181 AGAATATTTTAAAGTCTTATGATTAATGATTTAGTGTCTGATTTAAAAAAA 3240
 3241 AAA 3243
 3241 AAA 3243
 3241 AAA 3243
 3241 AAA 3243
 RESULT 2
 AF307339 3138 bp mRNA linear PRI 13-FEB-2001
 LOCUS AF307339
 DEFINITION Homo sapiens B aggressive lymphoma short isoform (BAL) mRNA,
 complete cds.
 ACCESSION AF307339
 VERSION AF307339.1 GI:12751140
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 (bases 1 to 3138)
 Aguiar,R.C., Yakushiji,Y., Kharbanda,S., Salgia,R., Fletcher,J.A.
 and Shipp,M.A.
 BAL is a novel risk-related gene in diffuse large B-cell lymphomas
 that enhances cellular migration
 Blood 96 (13), 4328-4334 (2000)
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 1110709
 2 (bases 1 to 3138)
 Aguiar,R.C.T., Yakushiji,Y. and Shipp,M.A.
 Direct Submission
 Submitted (21-SEP-2000) Adult Oncology, Dana Farber Cancer
 Institute, 44 Binney Street, Boston, MA 02115, USA

FEATURES
source
Location/Qualifiers
1. .3138
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3q13-q21"
1. .3138
/gene="BAL"
229. .2688
/gene="BAL"
/note="overexpressed in high risk diffuse large B-cell lymphoma"
/codon_start=1
/product="B aggressive lymphoma short isoform"
/protein_id="AAK07559.1"
/db_xref="GI:12751141"
/translation="MDPSWVAGAAAYNEKSENGALGSENYSMOPIRHNHDKILKNNR
OUCVUNKNFGCISTLVSPVQSGNSKSLQVFRKMLTPRIELSWKDLTHAVDAVN
AANEDLLHGGLALALVKAQGEIIESEKQFARVAKVAGELIAYGAGRLPKQIILH
AVGRMWMENDKQCTGKLRQALIVSLINLYIKTHIKTVAIPALISGIEOFPLNCTK
TIVETIRVSLQGRPMNSLKEIHLVSNEDPTVAFAKASEEILKSELGOETPSNA
MYNNLTLQVGHIEQMOTADYIVNSVNDHDIYGVVASKIIOQAGVEKMSIELATKA
KOPORSQILYLRKFNLFCKYIYHVMHSEPRPOLIKAMRECLKEIQNTSISF
PALGTENMEIKETAEILFDEVLTFAKHVHQDLTKFVLEPTDLEIKAESSEMAK
RSKMLSLNNYVSPQSTREKRENGLEARNLDMQENEMCEAHAWQRIISLQNH
HIENNHIILYLRKEHDIISLOKTSVSITELISGRTLEIEGARADLIEVNMIE
DMICKVOEEMARKKERGLRSLGOMTIOOKTODEKEMNIIFLKCPVPPLOLDOK
OFKSGLOYLKYERKIDNEVLMAAFORKKMEKILRQVSHSLFOQVYRQCNVYCR
VGFQNRSTPCDPKRYAGIYFTKNEKLNIAEKAKISAARKLIYFEAEVLTGFCQGH
PLNIVPPUSPGALDGHSDVNDVSSPEFVIFSGMOALPQYIMLTCTOEYVOSDYSS
GPMRPFQHPMRFAFSQPD"

BASE COUNT 962 a 622 c 733 g 821 t
ORIGIN

Query Match 93.2%; Score 3021.4; DB 9; Length 3138;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 3137; Conservative 0; Mismatches 1; Indels 105; Gaps 1;

QY 1 GGGCTTCGTTCTGCTGGTCTGCTGACGTCGCCGCCGCCGAGACTTGAAGTCT 60
DB 1 GGGCTTCGTTCTGCTGGTCTGCTGACGTCGCCGCCGCCGAGACTTGAAGTCT 60
QY 61 GGAAGTAGCTCTCCAGCTTCCTGTAATGCGGGGCGGAGCTTGACACCGCAGAGGA 120
DB 61 GGAAGTAGCTCTCCAGCTTCCTGTAATGCGGGGCGGAGCTTGACACCGCAGAGGA 120
QY 121 GCGGGGACGGCGCGGCGGAGAGTGCGCCACCATATCTGGAACCTATGCTTGA 180
DB 121 GCGGGGACGGCGCGGCGGAGAGTGCGCCACCATATCTGGAACCTATGCTTGA 180
QY 181 AGCGCAAAAGGGAATTAACATTTAAAGACTCCCGGGGAGCTGGAGATGGAATTTCC 240
DB 181 AGCGCAAAAGGGAATTAACATTTAAAGACTCCCGGGGAGCTGGAGATGGAATTTCC 240
QY 241 ATGGTGCGCGGAGCAGAGCTTCAATGAATAATCAGATGAGTTACCTGCTCTCATC 300
DB 241 ATGGTGCGCGGAGCAGAGCTTCAATGAATAATCAGATGAGTTACCTGCTCTCATC 300
QY 301 TTGTTTCAGAAAGCTTTTGTCTCAGATCTTTCCTCAGTGGAGAAAGGGAATACAGAA 360
DB 274 ----- 273
QY 361 TGTCCTCCCTACAAAGTCTCAGAGACTGTCCTTGGAGAAACATTAAGTGGCAATTT 420
DB 274 -----TAGAGACTGTCCTTGGAGAAACATTAAGTGGCAATTT 315
QY 421 CCCATTAAACCAATGACTCTCAAAATTTTAAAAAATTAATAGCGCTGAGTGTGAAGTC 480
DB 316 CCCATTAAACCAATGACTCTCAAAATTTTAAAAAATTAATAGCGCTGAGTGTGAAGTC 375
QY 481 CTCACAGATTAAGTTGGCTGTATCTCAACCTGCTCTTCCAGTTGAGGAAGCAGACG 540
DB 376 CTCACAGATTAAGTTGGCTGTATCTCAACCTGCTCTTCCAGTTGAGGAAGCAGACG 435

QY 541 AAATCTCTGCAAGTGTTCAGAAAAATGCTGACTCTAGATAGAGTATCAGTCTGGAAA 600
DB 436 AAATCTCTGCAAGTGTTCAGAAAAATGCTGACTCTAGATAGAGTATCAGTCTGGAAA 495
QY 601 GATGACCTCACACACATGCTGTTGATGCTGTGTGTAATCAGCAATGAAGATCTTCTG 660
DB 496 GATGACCTCACACACATGCTGTTGATGCTGTGTGTAATCAGCAATGAAGATCTTCTG 555
QY 661 CATGGGGAGGCGCTGGCGCTGGCGCTGTAAAGCTGTGGATTTGAAATCCAAAGAGAG 720
DB 556 CATGGGGAGGCGCTGGCGCTGGCGCTGTAAAGCTGTGGATTTGAAATCCAAAGAGAG 615
QY 721 AGCAAAAGTTTGTTCAGATATGTAAGAGTGCACCTGGTGAAGTGTGCACGGA 780
DB 616 AGCAAAAGTTTGTTCAGATATGTAAGAGTGCACCTGGTGAAGTGTGCACGGA 675
QY 781 GCAAGGAGGCTTCCCTGCAAAACAGATCATCATGCTGTTGGCGCTCGGTGGAATGG 840
DB 676 GCAAGGAGGCTTCCCTGCAAAACAGATCATCATGCTGTTGGCGCTCGGTGGAATGG 735
QY 841 GATTAACAGGAGTACTGTAAGAGCTGCAGAGGCGCATTTAGTATTTGAAATATGTC 900
DB 736 GATTAACAGGAGTACTGTAAGAGCTGCAGAGGCGCATTTAGTATTTGAAATATGTC 795
QY 901 ATCTATAAATATCTCATTAAGACAGTAGCAATTCCAGCTTGAAGCTTGGGATTTT 960
DB 796 ATCTATAAATATCTCATTAAGACAGTAGCAATTCCAGCTTGAAGCTTGGGATTTT 855
QY 961 CAGTTCCTCTGAAATTTGTGTACAAAGACTATGTAGAGACTATCCGGGTTAGTTGCAA 1020
DB 856 CAGTTCCTCTGAAATTTGTGTACAAAGACTATGTAGAGACTATCCGGGTTAGTTGCAA 915
QY 1021 GGAAGCCATGATGATATTTGAAGAAATTCACCTGCTGACCAATGAGACCCCTACT 1080
DB 916 GGAAGCCATGATGATATTTGAAGAAATTCACCTGCTGACCAATGAGACCCCTACT 975
QY 1081 GTTGCTGCTCTTAAAGCTGCTTCAAGTATCATCTAGAGAAAGTGAAGTGGGAAAGAA 1140
DB 976 GTTGCTGCTCTTAAAGCTGCTTCAAGTATCATCTAGAGAAAGTGAAGTGGGAAAGAA 1035
QY 1141 ACCACCCCTCTTCAATGCAATGAGTGTGTAACCACTGACCTCCAGATTTGCCAGGCG 1200
DB 1036 ACCACCCCTCTTCAATGCAATGAGTGTGTAACCACTGACCTCCAGATTTGCCAGGCG 1095
QY 1201 CACATTGAATGGCAGAGCGCAGATGTAATTTCTTAATCCCAATGATATTTACA 1260
DB 1096 CACATTGAATGGCAGAGCGCAGATGTAATTTCTTAATCCCAATGATATTTACA 1155
QY 1261 GTTGACCTGTGGCAAAAGTCAATTCACAAACAGCAGAGATGAAATGGAATTTT 1320
DB 1156 GTTGACCTGTGGCAAAAGTCAATTCACAAACAGCAGAGATGAAATGGAATTTT 1215
QY 1321 CTTGCCCAAAAGGCTAAACAGTTTCAACGGTCCAGTTGCTGTCACAAAAGATTT 1380
DB 1216 CTTGCCCAAAAGGCTAAACAGTTTCAACGGTCCAGTTGCTGTCACAAAAGATTT 1275
QY 1381 AACTGTGTTCTGTAATATATATACATGCTGTGGATTCAGAAATTTCTTAACCTCAG 1440
DB 1276 AACTGTGTTCTGTAATATATATACATGCTGTGGATTCAGAAATTTCTTAACCTCAG 1335
QY 1441 ATATTAACATGCAATGAAGAGTGTGTAAGAAATGCAATGAGCAAAATATTAACCTTC 1500
DB 1336 ATATTAACATGCAATGAAGAGTGTGTAAGAAATGCAATGAGCAAAATATTAACCTTC 1395
QY 1501 ATTTCTCTTTCCTGCTGAGCTGGAACATGAGAAATGAAGAGAAACAGCAGAGAG 1560
DB 1396 ATTTCTCTTTCCTGCTGAGCTGGAACATGAGAAATGAAGAGAAACAGCAGAGAG 1455
QY 1561 ATTTGTTTGAATGAAGTTTAAACATTTGCCAAGACCATGTAAACACCAAGTTAACTGTA 1620
DB 1456 ATTTGTTTGAATGAAGTTTAAACATTTGCCAAGACCATGTAAACACCAAGTTAACTGTA 1515

	Db	2596	TATGTACAGTCAACAAGATTACTCATCAGACCACAAAGAACCCTTTGGCACACATCCCTGG	2655
	QY	2761	AGGGATTTCGCAAGGCGAGGCCCTGGTGAATTAATCTCAATCATATTTTAAAGAGCTGTAT	2820
	Db	2656	AAGGATTCGCGAAGTGCGAGGCCCTGTGTATTAATCTCAATCATATTTTAAAGAGCTGTAT	2715
	QY	2821	GGCCTTACCTTGGGTGGAACATAACCAATATATGACCATCGATGCGTCAAAAGTAGGCTTGA	2880
	Db	2716	GGCCCTTACCTTGGGTGGAACATAACCAATATATGACCATCGATGCGTCAAAAGTAGGCTTGA	2775
	QY	2881	ATATATCCCATTGGGTATCTGTATGAGACTGACTGGGTATATGAAAAGCACTAGCCACATAC	2940
	Db	2776	ATATATCCCATTGGGTATCTGTATGAGACTGACTGGGTATATGAAAAGCACTAGCCACATAC	2835
	QY	2941	TAGCATCTTAGAGCCTTATATCTGTATATGTCCTTGGGGTGGGTAGTATACCAA	3000
	Db	2836	TAGCATCTTAGAGCCTTATATCTGTATATGTCCTTGGGGTGGGTAGTATACCAA	2895
	QY	3001	TGAAACACTTTCAGAGACCTCCCTCCCTGTCAGATGTCCTTAAATCTCTTACTAGAG	3060
	Db	2896	TGAAACACTTTCAGAGACCTCCCTCCCTGTCAGATGTCCTTAAATCTCTTACTAGAG	2955
	QY	3061	GAGATAAATATTTTGCATATATATGAGAAATTTTTCTGTATATATACCGAGCCTTTAT	3120
	Db	2956	GAGATAAATATTTTGCATATATATGAGAAATTTTTCTGTATATATACCGAGCCTTTAT	3015
	QY	3121	TTTCTTAAATATGATATATATATATATATATTTTGGATATACAGATGATTTTACATTTCCAG	3180
	Db	3016	TTTCTTAAATATGATATATATATATATATATTTTGGATATACAGATGATTTTACATTTCCAG	3075
	QY	3181	AGAAATATTTATAAAGGCTTATAGTATGAAAAATTAATCATCTTGTGTGATTTAAAAAAA	3240
	Db	3076	AGAAATATTTATAAAGGCTTATAGTATGAAAAATTAATCATCTTGTGTGATTTAAAAAAA	3135
	QY	3241	AAA 3243	
	Db	3136	AAA 3138	
RESULT 3				
LOCUS	AX399212			
DEFINITION	Sequence 1 from Patent WO0196399.			
ACCESSION	AX399212			
VERSION	AX399212.1			
KEYWORDS	GI:21261546			
SOURCE	.			
ORGANISM	human.			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE				
AUTHORS	Peyman,J.A., da Silva,A. and Hockman,P.			
TITLE	Novel interferon-induced tetraspan protein and nucleic acids encoding same			
JOURNAL	Patent: WO 0196399-A 1 20-DEC-2001			
FEATURES	Curegen Corporation (US); BIOGEN, INC. (US)			
	location/Qualifiers			
	1..3016			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	123..2582			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAD32679.1"			
	/db_xref="gi:21261547"			
	/translation="MDFSMVAAGAAAYNEKSETGALGENYSWQIPIINHDEFKILKNRER			
	OLCEVSLAKFGICSTLVSPOEGNNKSLOVERKKMLPRIELSYVKDGLTHAVDAVYN			
	AANDLLHGGLALAIKAGSEFOESKGFQFVARVGCVSGEIAIVTAGSLPKOIIH			
	AVGRWMHEDKOGCTGKLORAIRYSILNVYIKRNHTVALPALSSIRPOLPLNCRK			
	TIIVEITRSLEKGRPMKSNLKEIHLSVNEDPVAFAKASELTIGKSELGDPTSFNA			
	KVVNNVLIOVOGHIEWRTAPDIVANSNPDIIVPAKSLDOAGVMKSEFLATNAA			
	KOPRSOLVIETLKGFNFCKKIYHVLMHSERPQLIHKAMKECLECIQNITSISF			
CDS				

PALGTGMEIKKETAEILFDEVLTFAKDHVKHOLJTVKLVFVLEPDLLEYKAFSSBMAK
RSKMLSLNNYSVPOSTREKRENEGLEARSPAIINLMGEVEYEHAMJORIILSLONH
HIENNHLIYLRKEDHILISOLOKTSVSIPIEISPRELEIEBARDLEEVANIE
DMICKVOEMARKREGLMRISOWTIOOKDMKNNIIELEKPPVPOEILDOOK
OPEKGLQVLYEKIDNEVILMAAPORKKMMEKIRHROPVSRILEYEQUPYOFQNVYCR
VGFQWSTPCDPEKIGAGITFTKLNKINAEKAKKISADKRLIYEAEVLGFEQGH
PLNIVPLSLPGALDGHDSVVDNVSPTFVILFSQMOAIPOLIMTCTDEYVOSODYSS
GPMRFFAHPHMFSSGSPVD*
BASE COUNT 932 a 589 c 696 g 799 t
ORIGIN

Query Match 87.7%; Score 2842.6; DB 6; Length 3016;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 2963; Conservative 0; Mismatches 9; Indels 105; Gaps 1;

QY 151 CAPATCTGGAACTACAGTCTATGCTTGAAGCGCAAAAGGGAATTAACATTAAAGACT 210
DB 45 CAGGCTGGAACTACAGTCTATGCTTGAAGCGCAAAAGGGAATTAACATTAAAGACT 104
QY 211 CCCCCGGGACCTGAGAGATGACCTTTTCCATGCTGGCCGAGACAGCAGCTTACAATGAA 270
DB 105 CCCCCGGGACCTGAGAGATGACCTTTTCCATGCTGGCCGAGACAGCAGCTTACAATGAA 164
QY 271 AATACAGTAGATTACCTCGCTCTCAGTCTGTTTCAAGAACTTTTGCTCAGATCTTT 330
DB 165 AAA----- 167
QY 331 CCTCAGTGAGAAAGGGAATACAGAAAGATGTCCTCCCTACAAAGTGTCCAGAGACTGGT 390
DB 168 ----- -TCAGAGACTGGT 179
QY 391 GCTCTTGGAGAAAACATATAGTTGGCAAAATTCOCATTAACACATGACTTCAAAATTTTA 450
DB 180 GCTCTTGGAGAAAACATATAGTTGGCAAAATTCOCATTAACACATGACTTCAAAATTTTA 239
QY 451 AAAAATAATGAGACGCTGCTGTGTGAAGTCCACAGATAAGTTTGGCTGTATCTTACC 510
DB 240 AAAAATAATGAGACGCTGCTGTGTGAAGTCCACAGATAAGTTTGGCTGTATCTTACC 299
QY 511 CTGTGCTCTCCAGTTCCAGGAGGCAACAGCAAAATCTCTCAAGTGTTCAGAAAAATGCTG 570
DB 300 CTGTGCTCTCCAGTTCCAGGAGGCAACAGCAAAATCTCTCAAGTGTTCAGAAAAATGCTG 359
QY 571 ACTCTAGAGATAGATTATCAGTCTGTGGAAGATGACCTCACACATGCTGTGTGATGCT 630
DB 360 ACTCTAGAGATAGATTATCAGTCTGTGGAAGATGACCTCACACATGCTGTGTGATGCT 419
QY 631 GTGTGTAATGACGCAATGAAGATCTTCTGCATGGGGGAGGCTGGCCCTCGCTGTGTA 690
DB 420 GTGTGTAATGACGCAATGAAGATCTTCTGCATGGGGGAGGCTGGCCCTCGCTGTGTA 479
QY 691 AAAGGCTGGATTTGAATCCAGAAGAGACAACAGCTTTGTTCCAGATATGTGTA 750
DB 480 AAAGGCTGGATTTGAATCCAGAAGAGACAACAGCTTTGTTCCAGATATGTGTA 539
QY 751 GTGTACAGTGTGAGATAGCTGTACAGGAGGAGGAGGCTTCCCTGCAACAGATCATC 810
DB 540 GTGTACAGTGTGAGATAGCTGTACAGGAGGAGGAGGCTTCCCTGCAACAGATCATC 599
QY 811 CATGCTGTGGGCTCTGGGTGATGGAATGGGATAAACAGGAGTGTACTGGAAGCTGCGAG 870
DB 600 CATGCTGTGGGCTCTGGGTGATGGAATGGGATAAACAGGAGTGTACTGGAAGCTGCGAG 659
QY 871 AAGGCCATGTAGATTTGTGAATTTGTCAATCTATAAAAATACACATTAAGACAGTA 930
DB 660 AAGGCCATGTAGATTTGTGAATTTGTCAATCTATAAAAATACACATTAAGACAGTA 719
QY 931 GCAATTTCCAGCCTTGAGCTCTGGGATTTTTCAGTTCCCTCTGTAATTTGTGTAAGACT 990
DB 720 GCAATTTCCAGCCTTGAGCTCTGGGATTTTTCAGTTCCCTCTGTAATTTGTGTAAGACT 779
QY 991 ATTGTAGACATATCCGGGTTAGTTTGAAGGAAACCAATGATGATTAATTTGAAAGAA 1050

DB 780 ATTGTAGACATATCCGGGTTAGTTTGCAGGGAGCAACATGATGATTAATTTGAAAGAA 839
QY 1051 ATTACACCTGGTGAGCAATGAGACCCCTACTGTTGCTGCCCTTTAAAGCTGCTCAGAAATTC 1110
DB 840 ATTACACCTGGTGAGCAATGAGACCCCTACTGTTGCTGCCCTTTAAAGCTGCTCAGAAATTC 899
QY 1111 ATTCCTAGGAAAGTAGAGTGGGAGACAAGAACCCCTCTTCTTCAATGCAATGTCGTG 1170
DB 900 ATTCCTAGGAAAGTAGAGTGGGAGACAAGAACCCCTCTTCTTCAATGCAATGTCGTG 959
QY 1171 AACACCTGACCTCCAGATTTGTCCAGGGCCACATGTAATGGCAGACGACATGTATTT 1230
DB 960 AACACCTGACCTCCAGATTTGTCCAGGGCCACATGTAATGGGAGGAGGAGTGTAAATTT 1019
QY 1231 GTTAAATTTGTAAACCCACATGATTTATAGATTGAGCTGTGGCAAGTCAATTTACAA 1290
DB 1020 GTTAAATTTGTAAACCCACATGATTTATAGATTGAGCTGTGGCAAGTCAATTTACAA 1079
QY 1291 CAAGCAGAGAGTGAATGGAATTCGAATTTCTTGCCACAAGGCTTAACAGTTTCAAGG 1350
DB 1080 CAAGCAGAGAGTGAATGGAATTCGAATTTCTTGCCACAAGGCTTAACAGTTTCAAGG 1139
QY 1351 TCCCATGTTGTAAGTGTGTCACAAAAGATTTAACTGTCTGTAAATATATATACATGTA 1410
DB 1140 TCCCATGTTGTAAGTGTGTCACAAAAGATTTAACTGTCTGTAAATATATATACATGTA 1199
QY 1411 CTGTGCAATTCAGAAATTTCTTAAACCTCAGATTTTAAACATGCAATGAAAGAGTGTG 1470
DB 1200 CTGTGCAATTCAGAAATTTCTTAAACCTCAGATTTTAAACATGCAATGAAAGAGTGTG 1259
QY 1471 GAAAAATGATGAGCAAAATATACCTTCATTTCTGTGCTGCTGGAGTGTGAAC 1530
DB 1260 GAAAAATGATGAGCAAAATATACCTTCATTTCTGTGCTGCTGGAGTGTGAAC 1319
QY 1531 ATGGAATTAAGAGGAACAGACAGAGATTTGTTGATGAATTTAACTTAATTTGGC 1590
DB 1320 ATGGAATTAAGAGGAACAGACAGAGATTTGTTGATGAATTTAACTTAATTTGGC 1379
QY 1591 AAAGACCATGTAAACACACAGTAACTGTAAATTTGTATGATCTTTCACACAGATTTGGAG 1650
DB 1380 AAAGACCATGTAAACACACAGTAACTGTAAATTTGTATGATCTTTCACACAGATTTGGAG 1439
QY 1651 ATATATAAAGCTTTTCACTGTTCTGAATGCAAAAGAGTCCAAAGTGTGATTTGAACAT 1710
DB 1440 ATATATAAAGCTTTTCACTGTTCTGAATGCAAAAGAGTCCAAAGTGTGATTTGAACAT 1499
QY 1711 TACAGTGTCCCCAGTCAACACAGAGAGAGAAAGAAAGGAGGCTTAAGCTAATGCT 1770
DB 1500 TACAGTGTCCCCAGTCAACACAGAGAGAGAAAGAAAGGAGGCTTAAGCTAATGCT 1559
QY 1771 CCTGCCATCAATCTGATGGGATTTCAACGTGAAGAGATGATGAGGCCACAGCATGATC 1830
DB 1560 CCTGCCATCAATCTGATGGGATTTCAACGTGAAGAGATGATGAGGCCACAGCATGATC 1619
QY 1831 CAAAGAACTCTGAGTCTCCAGAACCAACACATCATTTGAGAATATATCTGTACCTTT 1890
DB 1620 CAAAGAACTCTGAGTCTCCAGAACCAACACATCATTTGAGAATATATCTGTACCTTT 1679
QY 1891 GGGAGAAAGAACATGACATTTTGTCTCAGCTTCAGAAAATTTCAAGTGTCTCCATCACA 1950
DB 1680 GGGAGAAAGAACATGACATTTTGTCTCAGCTTCAGAAAATTTCAAGTGTCTCCATCACA 1739
QY 1951 GAAATTTATCAGCCAGGAGAGACAGATTTAGAGATTTGAAGAGCCGGGCTCAGCTCATTT 2010
DB 1740 GAAATTTATCAGCCAGGAGAGACAGATTTAGAGATTTGAAGAGCCGGGCTCAGCTCATTT 1799
QY 2011 GAGGTGGTTATGACATTTGAAGATATGCTTTGTAAAGTACAGAGGAATTTGCAAGGAA 2070
DB 1800 GAGGTGGTTATGACATTTGAAGATATGCTTTGTAAAGTACAGAGGAATTTGCAAGGAA 1859
QY 2071 AAGGAGCGAGGCTTTGGGCTGCTTTAGACAGTGGACATTTTCAGACACAAAAACCCAA 2130
DB 1860 AAGGAGCGAGGCTTTGGGCTGCTTTAGACAGTGGACATTTTCAGACACAAAAACCCAA 1919

QY 2131 GAGGAATGAAGAAATATCATATTTCTGAAATGTCCTGCTCCAGCTCAAGAGCTT 2190
|||||
Db 1920 GAGGAATGAAGAAATATCATATTTCTGAAATGTCCTGCTCCAGCTCAAGAGCTT 1919
QY 2191 CTAGATCAAAAAGAAACAGTTGAAAAATGTTGGAGGTTCTAAAGGTGGAGAGATA 2250
|||||
Db 1980 CTAGATCAAAAAGAAACAGTTGAAAAATGTTGGAGGTTCTAAAGGTGGAGAGATA 2039
QY 2251 GACAATGAGTCTCTATAGGCTGCTTCAAGAAAGAAAGAAATGATGGAAGAAAACTG 2310
|||||
Db 2040 GACAATGAGTCTCTATAGGCTGCTTCAAGAAAGAAAGAAATGATGGAAGAAAACTG 2099
QY 2311 CACAGGCAACCTGTGAGCCATGAGGCTGTTGAGCAAGTCCCATCCAGTCTGCAAGTGTG 2370
|||||
Db 2100 CACAGGCAACCTGTGAGCCATGAGGCTGTTGAGCAAGTCCCATCCAGTCTGCAAGTGTG 2159
QY 2371 GTATGCAAGTGTGCTTCAAGAAATGATCTGACACCTTGGCATGCCAAATACGAGAGCT 2430
|||||
Db 2160 GTATGCAAGTGTGCTTCAAGAAATGATCTGACACCTTGGCATGCCAAATACGAGAGCT 2219
QY 2431 GGCATATCTTCAACCAAGAACTTCAAAAACCTGGCAGAGAAAGCCCAAGAAATCTGCT 2490
|||||
Db 2220 GGCATATCTTCAACCAAGAACTTCAAAAACCTGGCAGAGAAAGCCCAAGAAATCTGCT 2279
QY 2491 GCAAGATAGCGATCTATGTTGTTGAGGCTGAGTACTCAACAGGCTCTTCTGCCAGGGA 2550
|||||
Db 2280 GCAAGATAGCGATCTATGTTGTTGAGGCTGAGTACTCAACAGGCTCTTCTGCCAGGGA 2339
QY 2551 CATCCGTTAAATATGTTCCCCACACACTAGATCTTGAGCTATAGATGTCATGAGAGT 2610
|||||
Db 2340 CATCCGTTAAATATGTTCCCCACACTAGATCTTGAGCTATAGATGTCATGAGAGT 2399
QY 2611 GTGTTGACAAATGCTCTCCAGCCCTGAAACCTTGTATTTTACTGTCATGACGCTATA 2670
|||||
Db 2400 GTGTTGACAAATGCTCTCCAGCCCTGAAACCTTGTATTTTACTGTCATGACGCTATA 2459
QY 2671 CCTCAGATTTTGGAGCATGACACCCAGAAATATGTACAGTACACAAATCTCTCATCGA 2730
|||||
Db 2460 CCTCAGATTTTGGAGCATGACACCCAGAAATATGTACAGTACACAAATCTCTCATCGA 2519
QY 2731 CCAATGAGACCTTGGCAGACATCTTGGAGGGATTTGCCAAATGGCAGCCCTGTGAT 2790
|||||
Db 2520 CCAATGAGACCTTGGCAGACATCTTGGAGGGATTTGCCAAATGGCAGCCCTGTGAT 2579
QY 2791 TAACTCTAATCATCTTTTAAACAGCTGTATGCTTACCTTGGGTGAACCTAACCAATTA 2850
|||||
Db 2580 TAACTCTAATCATCTTTTAAACAGCTGTATGCTTACCTTGGGTGAACCTAACCAATTA 2639
QY 2851 TGACCATGAGTGGCTCAAGAGTGGCTTGAATATATCCCATGGGTATTCGTATGAGACTG 2910
|||||
Db 2640 TGACCATGAGTGGCTCAAGAGTGGCTTGAATATATCCCATGGGTATTCGTATGAGACTG 2699
QY 2911 ACTGGGTTATTTGAAGAGACTAGCCATACATAGCATCTTATGCTTATCTCTTAT 2970
|||||
Db 2700 ACTGGGTTATTTGAAGAGACTAGCCATACATAGCATCTTATGCTTATCTCTTAT 2759
QY 2971 GCTCTGGGTTGGGGTGGAGTACCAATGAACACTTTCAGAGCTTCTCTCTCT 3030
|||||
Db 2760 GCTCTGGGTTGGGGTGGAGTACCAATGAACACTTTCAGAGCTTCTCTCTCTCT 2819
QY 3031 GCACTTCTTCTTATCTCTTCTTCTAGAGAGATTAATTTTGGCATATATAGAGAA 3090
|||||
Db 2820 GCACTTCTTCTTATCTCTTCTTCTAGAGAGATTAATTTTGGCATATATAGAGAA 2879
QY 3091 TTTTCTAGATATATAGCAGGCTTTTATTTCTAAATCATGATATGTAATTAATGCTT 3150
|||||
Db 2880 TTTTCTAGATATATAGCAGGCTTTTATTTCTAAATCATGATATGTAATTAATGCTT 2939
QY 3151 AGGATTAACAGAAATGATTTTATGATTTTCCAGAGATATTTATTAAGTCTTATAGTATGAA 3210
|||||
Db 2940 AGGATTAACAGAAATGATTTTATGATTTTCCAGAGATATTTATTAAGTCTTATAGTATGAA 2999

QY 3211 ATAAATCATCTTGTCT 3227
|||||
Db 3000 ATAAATCATCTTGTCT 3016

RESULT 4
AX274946 2598 bp DNA linear PAT 29-OCT-2001
LOCUS
DEFINITION Sequence 211 from Patent WO0172777.
ACCESSION AX274946
VERSION AX274946.1 GI:16547578
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Hillman, J.L., Baughn, M.R., Yue, H., Lal, P., Lu, D.A., Patterson, C., Azimzai, Y., Bandman, O., Tang, Y.T., Mathur, P., Shah, P., Au-Young, J. and Reddy, R.
Transcription factors
Patent: WO 017277-A 211 04-OCT-2001;
Incyte Genomics, Inc. (US)
location/Qualifiers
1..2598
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 557394CB1"

BASE COUNT 823 a 491 c 587 g 697 t
ORIGIN

Query Match 78.4%; Score 2543.8; DB 6; Length 2598;
Best Local Similarity 99.98; Pred. No. 0;
Matches 2545; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 696 TGGTGATTTGAAATCCAGAAGAGAGCAACAGTTTGTTCGAGATATGTAAGTGTG 755
|||||
Db 52 TGGTGATTTGAAATCCAGAAGAGAGCAACAGTTTGTTCGAGATATGTAAGTGTG 111
QY 756 AGCTGTGAGATAGCTGTACAGGAGAGAGAGGCTTCCGCAACAGATCATCATGC 815
|||||
Db 112 AGCTGTGAGATAGCTGTACAGGAGAGAGAGGCTTCCGCAACAGATCATCATGC 171
QY 816 TGTGGGCTGTGGTGAATGGAATGGAATGAACAGGAGTGTCTGGAAGCTGCAGAGGCG 875
|||||
Db 172 TGTGGGCTGTGGTGAATGGAATGGAATGAACAGGAGTGTCTGGAAGCTGCAGAGGCG 231
QY 876 CATTGTAAGTATTTGTAATTTATGTCATCTATAAATATCTACATTAAGACATAGCAAT 935
|||||
Db 232 CATTGTAAGTATTTGTAATTTATGTCATCTATAAATATCTACATTAAGACATAGCAAT 291
QY 936 TCCAGCCTTGAGCTGTGGGATTTTTCAGTCCCTGTGATTTGTGTACAAAGACTATTGT 995
|||||
Db 292 TCCAGCCTTGAGCTGTGGGATTTTTCAGTCCCTGTGATTTGTGTACAAAGACTATTGT 351
QY 996 AGAGACATCCGGGTTAGTTTGCAGAGGAAGCCATGATGTAATTTGAAGAAATTTCA 1055
|||||
Db 352 AGAGACATCCGGGTTAGTTTGCAGAGGAAGCCATGATGTAATTTGAAGAAATTTCA 411
QY 1056 CCTGTGAGCAATAGAGACCTTACTGTTGTGCTGCTTTAAACCTCTCAGATTCATCCT 1115
|||||
Db 412 CCTGTGAGCAATAGAGACCTTACTGTTGTGCTGCTTTAAACCTCTCAGATTCATCCT 471
QY 1116 AGGAGAGAGTGAAGTGGAGCAAGAAACACCCCTTCTTCAATGCAATGTCGTGAACAA 1175
|||||
Db 472 AGGAGAGAGTGAAGTGGAGCAAGAAACACCCCTTCTTCAATGCAATGTCGTGAACAA 531
QY 1176 CCTGACCTCCAGATTGTCAGAGGCCACATTGAATGGACAGCGGAGAGTGAATTTGTTAA 1235
|||||
Db 532 CCTGACCTCCAGATTGTCAGAGGCCACATTGAATGGACAGCGGAGAGTGAATTTGTTAA 591
QY 1236 TTTCTTAACCCACATGATATTTACAGTTTGAAGCTGTGGCAAGTCAATTTCAACAAGC 1295
|||||

Db 592 TTCTTAAACCCACATGATATTACAGTTGGACCTTGGCAAAAGTAAATCTTCAACAAACG 651
Qy 1296 AGAGCTTGAATGAATGGGAATTTCTTGGCCACAAAGGCTAAACAGTTTCAACGGTCCCA 1355
Db 652 AGAGCTTGAATGAATGGGAATTTCTTGGCCACAAAGGCTAAACAGTTTCAACGGTCCCA 711
Qy 1356 GTTGGTACTGGTCACAAAAGCATTTAACTTGTCTGTAAATATATATACATGTACTGTG 1415
Db 712 GTTGGTACTGGTCACAAAAGCATTTAACTTGTCTGTAAATATATATACATGTACTGTG 771
Qy 1416 GCATTCAGAAATTTCTAAACCTCAGATATTAACATGCAATGAAGAGATGTTGGAAA 1475
Db 772 GCATTCAGAAATTTCTAAACCTCAGATATTAACATGCAATGAAGAGATGTTGGAAA 831
Qy 1476 ATGCATTGAGCAAAATATTAATCTTCATTTCTTCCCTTGGGACTGGAACATGGA 1535
Db 832 ATGCATTGAGCAAAATATTAATCTTCATTTCTTCCCTTGGGACTGGAACATGGA 891
Qy 1536 AATAAAGAAGAAACAGCAGCAGAGATTTTGTGTGAAGTTTAACTTTTCCCAAAGA 1595
Db 892 AATAAAGAAGAAACAGCAGCAGATTTTGTGTGAAGTTTAACTTTTCCCAAAGA 951
Qy 1596 CCATGTAAACACCGATTAACGTAAATTTGTGATCTTCCACAGATTTGGAGATATA 1655
Db 952 CCATGTAAACACCGATTAACGTAAATTTGTGATCTTCCACAGATTTGGAGATATA 1011
Qy 1656 TAAGCTTTTCACTTGTGAATATGCAAGAGTCCAAAGATGCTGAGTTTGACAAATTACAG 1715
Db 1012 TAAGCTTTTCACTTGTGAATATGCAAGAGTCCAAAGATGCTGAGTTTGACAAATTACAG 1071
Qy 1716 TGTCCCCAGTAAACCAAGAGAGAGAGAAAAGAAAATGGGCTTGAAGCTTACCTTCC 1775
Db 1072 TGTCCCCAGTAAACCAAGAGAGAGAGAAAAGAAAATGGGCTTGAAGCTTACCTTCC 1131
Qy 1776 CATCATCTGATGGGATTCACGTGGAAGAGATGATGAGGCCCGACGGATCCCAAG 1835
Db 1132 CATCATCTGATGGGATTCACGTGGAAGAGATGATGAGGCCCGACGGATCCCAAG 1191
Qy 1836 AATCTGATCTCCAGAAACACACACATCATGAGATATATCATATTTCTTACCTTGGAG 1895
Db 1192 AATCTGATCTCCAGAAACACACACATCATGAGATATATCATATTTCTTACCTTGGAG 1251
Qy 1896 AAGGATCATGACATTTTGTCTCAGCTTCAGAAAACCTTCACTGTCTCATACAGAAAT 1955
Db 1252 AAGGATCATGACATTTTGTCTCAGCTTCAGAAAACCTTCACTGTCTCATACAGAAAT 1311
Qy 1956 TATCAGCCAGAGACAGAGATGAGATGAGAGAGCCCGGGCTGACCTCATTTAGAT 2015
Db 1312 TATCAGCCAGAGAGACAGAGATGAGATGAGAGAGCCCGGGCTGACCTCATTTAGAT 1371
Qy 2016 GGTATGACATTTGAATATGCTTTGTAAATGATACAGAGAAATGCGAAGAAAAAGA 2075
Db 1372 GGTATGACATTTGAATATGCTTTGTAAATGATACAGAGAAATGCGAAGAAAAAGA 1431
Qy 2076 GCGAGGCTTTGGCGCTGTAGGACAGTGAATTCAGCAACAAAAAACCAGAGCA 2135
Db 1432 GCGAGGCTTTGGCGCTGTAGGACAGTGAATTCAGCAACAAAAAACCAGAGCA 1491
Qy 2136 AATGAAAGAAATATCATATTTCTGAATATGCTGTGCTCCCAACTCAAGACTTCTANA 2195
Db 1492 AATGAAAGAAATATCATATTTCTGAATATGCTGTGCTCCCAACTCAAGACTTCTANA 1551
Qy 2196 TCAAAAGAAACAGTTGAAAATGTGTTGAGGTTCTTAAAGGTGAGAGATAGACAA 2255
Db 1552 TCAAAAGAAACAGTTGAAAATGTGTTGAGGTTCTTAAAGGTGAGAGATAGACAA 1611
Qy 2256 TGAGTCTCTTATGCTGCTTTCAAGAAAGAAATGATGAGAAAGAAACTGCACAG 2315
Db 1612 TGAGTCTCTTATGCTGCTTTCAAGAAAGAAATGATGAGAAAGAAACTGCACAG 1671
Qy 2316 GCAACCTGTGACCAATAGGCTTTTCCAGCAATCCCATCCAGTCTTGCATATGTGTAG 2375
Db 1672 GCAACCTGTGACCAATAGGCTTTTCCAGCAATCCCATCCAGTCTTGCATATGTGTAG 1731

Qy 2376 CAGAGTTGGCTTTCAAAGATGTACTCGACACCTTGGCATCCAAAATACGAGCTGGCAT 2435
Db 1732 CAGAGTTGGCTTTCAAAGATGTACTCGACACCTTGGCATCCAAAATACGAGCTGGCAT 1791
Qy 2436 ATACCTTCACCAAGAACCTTCAAAAACCTGGCAGAGAAAGGCAAGAAATCTGTGCAG 2495
Db 1792 ATACCTTCACCAAGAACCTTCAAAAACCTGGCAGAGAAAGGCAAGAAATCTGTGCAG 1851
Qy 2496 TAAGCTGATCTATGTGTGAGGCTGAATGATCTCAGAGCTTCTTCTGCAGAGCATCC 2555
Db 1852 TAAGCTGATCTATGTGTGAGGCTGAATGATCTCAGAGCTTCTTCTGCAGAGCATCC 1911
Qy 2556 GTTAAATATTTGTTCCCCACACATGATCTGAGAGCTATGATGCTATGACATGTGCT 2615
Db 1912 GTTAAATATTTGTTCCCCACACATGATCTGAGAGCTATGATGCTATGACATGTGCT 1971
Qy 2616 TGACAAATGCTCCAGCCCTGAAACCTTGTATTTTATGTCATGACAGCATACCTCA 2675
Db 1972 TGACAAATGCTCCAGCCCTGAAACCTTGTATTTTATGTCATGACAGCATACCTCA 2031
Qy 2676 GTATTTGTGACATGACACCCAGAAATATGTACAGTCACAAAGATTAATCTCAGACCAAT 2735
Db 2032 GTATTTGTGACATGACACCCAGAAATATGTACAGTCACAAAGATTAATCTCAGACCAAT 2091
Qy 2736 GAGACCTTTTGACAGACATCTTTGAGGGGATTCGCAAGTGCAGCCCTGTTGATTAATC 2795
Db 2092 GAGACCTTTTGACAGACATCTTTGAGGGGATTCGCAAGTGCAGCCCTGTTGATTAATC 2151
Qy 2796 TCTACATCATTTTAAACAGCTGTATGCTTACCTTGGGTGAACATTAACCAATATGACC 2855
Db 2152 TCTACATCATTTTAAACAGCTGTATGCTTACCTTGGGTGAACATTAACCAATATGACC 2211
Qy 2856 ATCGATGCTCAAGAGTGGCTTGAATATATCCATGGGTATCTGTATGACTGACTGG 2915
Db 2212 ATCGATGCTCAAGAGTGGCTTGAATATATCCATGGGTATCTGTATGACTGACTGG 2271
Qy 2916 GTTATTTGAAGAGATAGCCACATCTAGCATCTTATGCTTATCTTATCTT 2975
Db 2272 GTTATTTGAAGAGATAGCCACATCTAGCATCTTATGCTTATCTTATCTT 2331
Qy 2976 GGGGTTGGGATGATGATACCAATGAAGACCTTCAGAGACTTCCCTCTTGCAGT 3035
Db 2332 GGGGTTGGGATGATGATACCAATGAAGACCTTCAGAGACTTCCCTCTTGCAGT 2391
Qy 3036 TGTCTTTTATCTCCTTCTAGAGAGATTAATTTTGCATATATGAAGAAATTTT 3095
Db 2392 TGTCTTTTATCTCCTTCTAGAGAGATTAATTTTGCATATATGAAGAAATTTT 2451
Qy 3096 CTAGTATATACGACAGGCTTTTATTTCTAAATGATGATAGTAAAAATCTTAGAT 3155
Db 2452 CTAGTATATACGACAGGCTTTTATTTCTAAATGATGATAGTAAAAATCTTAGAT 2511
Qy 3156 AACAGATGATTTTATGATTTTCCAGAGATATTAATGCTTTTATGATGATGATTA 3215
Db 2512 AACAGATGATTTTATGATTTTCCAGAGATATTAATGCTTTTATGATGATGATTA 2571
Qy 3216 TCATCTTGTCTGATTAATAAAAAA 3242
Db 2572 TCATCTTGTCTGATTAATAAAAAA 2598

RESULT 5
HSM803023 HSM803023 1442 bp mRNA linear PRI 20-MAR-2002
LOCUS HSM803023
DEFINITION Homo sapiens mRNA; cDNA DKFZp76111617 (from clone DKFZp76111617).
ACCESSION AL713679
VERSION AL713679.1 GI:19584378
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1442)
AUTHORS Ansoore,W., Winkner,U., Mewes,H.W., Well,B. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2002) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp761I1617) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.
FEATURES
source
Location/Oallifiers
1..1442
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="3"
/clone="DKFZp761I1617"
/issue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..1442
/gene="DKFZp761I1617"
<1..993
/gene="DKFZp761I1617"
/note="BAL"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD284379.1"
/db_xref="GI:19584379"
/translation="EMYEAHAWTORILSLIONHHIENNHLILYGRKEHDLISLOLKT
SSVITTEIISGRTELEIGARADLIEVWNIIDMLCKVDEMARKKERGLMSIGOM
TIOOKTODEKEMNIIFLKCPPTQELLDQKQEKGLVLEKEDNVLMAFG
RKKKMEKELHROVSHRLEFOOVYPCNVRGFORMYSTPOPKVGAIFYEKNL
KILAEKAKTISAADKLIVFEAFVLTGFRGCHPNTYPPRLSPADIGHSVDNVS
SPETVFIFSGQALIPQYIMTCTQEVOSQDYSGBMRFAQHPMRGAFSGSPVD"
polyA_signal
1410..1415
/gene="DKFZp761I1617"
1431
polyA_site
/gene="DKFZp761I1617"
BASE COUNT 457 a 276 c 317 g 392 t
ORIGIN
Query Match 44.4%; Score 1440.4; DB 9; Length 1442;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1441; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
polyA_signal
1410..1415
/gene="DKFZp761I1617"
1431
polyA_site
/gene="DKFZp761I1617"
BASE COUNT 457 a 276 c 317 g 392 t
ORIGIN

Db 301 CAGTGACTATTTCAGCAACAAAACCCAGAGCAAGTAAGAAATATCATATTCTCG 360
Qy 2161 AATGTCTGTCTCCCTCCCAACTCAAGAGCTTCTAGATCAAAAACAGTTTAAAAATGT 2220
Db 361 AATGTCTGTCTCCCTCCCAACTCAAGAGCTTCTAGATCAAAAACAGTTTAAAAATGT 420
Qy 2221 GATTTGACAGCTCTAAGGTGAGAAAGATGACATGAGAGTCTTATGCTCCCTTCA 2280
Db 421 GATTTGACAGCTCTAAGGTGAGAAAGATGACATGAGAGTCTTATGCTCCCTTCA 480
Qy 2281 AGAAGAGAGAAATATGAGAGAGAAACATGACAGCAACCTGTAGCCATAGCTGT 2340
Db 481 AGAAGAGAGAAATATGAGAGAGAAACATGACAGCAACCTGTAGCCATAGCTGT 540
Qy 2341 CAGCAAGTCCCATACAGCTTCTGCAATGTGTATGACAGAGTTGGCTTCAAGAAATGAC 2400
Db 541 CAGCAAGTCCCATACAGCTTCTGCAATGTGTATGACAGAGTTGGCTTCAAGAAATGAC 600
Qy 2401 TCGACACCTTGAGATCCAAATATGAGAGCTGATATACCTGACCAAGAACTCAAAAAC 2460
Db 601 TCGACACCTTGAGATCCAAATATGAGAGCTGATATACCTGACCAAGAACTCAAAAAC 660
Qy 2461 CTGGCAGAGAGAGCCAGAGAAATCTCTGCTGACAGATAGCTGATCTATGTGTAGAGCT 2520
Db 661 CTGGCAGAGAGAGCCAGAGAAATCTCTGCTGACAGATAGCTGATCTATGTGTAGAGCT 720
Qy 2521 GAAAGTACTCAGAGCTTCTTCTGCCAGAGACATCCGTTAAATATGTTCCTCCACACCTG 2580
Db 721 GAAAGTACTCAGAGCTTCTTCTGCCAGAGACATCCGTTAAATATGTTCCTCCACACCTG 780
Qy 2581 ACTCTGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
Db 781 ACTCTGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Qy 2641 TTTGTTATTTTATGAGCATGAGAGCTATACCTGATATTTGTGACATGACACCAAGAA 2700
Db 841 TTTGTTATTTTATGAGCATGAGAGCTATACCTGATATTTGTGACATGACACCAAGAA 900
Qy 2701 TATGTACATGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
Db 901 TATGTACATGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Qy 2761 AGGGGATGCAAGTGGAGCCCTGTTGATGATGATGATGATGATGATGATGATGATGAT 2820
Db 961 AGGGGATGCAAGTGGAGCCCTGTTGATGATGATGATGATGATGATGATGATGATGAT 1020
Qy 2821 GGCCTTACCTGGTGAATCAACCAATATGACATGATGATGATGATGATGATGATGATGAT 2880
Db 1021 GGCCTTACCTGGTGAATCAACCAATATGACATGATGATGATGATGATGATGATGATGAT 1080
Qy 2881 ATATATCCATGGTGTATCTGTATGATGATGATGATGATGATGATGATGATGATGAT 2940
Db 1081 ATATATCCATGGTGTATCTGTATGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy 2941 TAGCATCTTAGAGCTTATCTGTCTTATGATGATGATGATGATGATGATGATGATGAT 3000
Db 1141 TAGCATCTTAGAGCTTATCTGTCTTATGATGATGATGATGATGATGATGATGATGAT 1200
Qy 3001 TGAACACTTTCAGAGACCTTCTCTCTGAGAGTGTCTTAAATCTCTCTTACTAGAG 3060
Db 1201 TGAACACTTTCAGAGACCTTCTCTCTGAGAGTGTCTTAAATCTCTCTTACTAGAG 1260
Qy 3061 GAGATTAATATTTGATATTAATGAAGAAATTTTCTACTATATTAACGAGGCTTTTAT 3120
Db 1261 GAGATTAATATTTGATATTAATGAAGAAATTTTCTACTATATTAACGAGGCTTTTAT 1320
Qy 3121 TTTCTAAATGATGATAGATATTAATGATAGATTAATGATAGATTAATGATAGATTAAT 3180
Db 1321 TTTCTAAATGATGATAGATATTAATGATAGATTAATGATAGATTAATGATAGATTAAT 1380
Qy 3181 AGAATATTAATTAAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Db 1381 AGAATATTAATTAAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1440

QY 3241 AA 3242
Db 1441 AA 1442

RESULT 6
HSM804240
LOCUS HSM804240 1388 bp mRNA linear PRI 10-JUL-2002
DEFINITION Homo sapiens mRNA: cDNA DKFzp66B0810 (from clone DKFzp66B0810).
ACCESSION AL832929
VERSION AL832929.1 GI:21733516
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.
1 (bases 1 to 1388)
Ansoerge,W., Winkner,U., Mewes,H.W., Weil,B. and Wiemann,S.
REFERENCE Direct Submission
AUTHORS Submitted (09-JUL-2002) 1, D-85764 Neuberberg, GERMANY
TITLE Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de,
JOURNAL Heidelberg/Germany) within the CDNA sequencing consortium of the
COMMENT This clone (DKFzp66B0810) is available at the RZPD in Berlin.
This clone (DKFzp66B0810) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clonerzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.
Location/Qualifiers
source 1..1388
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFzp66B0810"
/tissue_type="stomach"
/clone_lib="666 (synonym: hsto2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
polya_signal 1342..1347
polya_site 1360
BASE COUNT 450 a 257 c 302 g 379 t
ORIGIN

Query Match 42.3%; Score 1371.8; DB 9; Length 1388;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1373; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1869 GAATTAATCATATTTCTGACCTGGGAGAGAAAGACATGACATTTTGTCTCAGCTTCAGAA 1928
Db 1 GAATTAATCATATTTCTGACCTGGGAGAGAAAGACATGACATTTTGTCTCAGCTTCAGAA 60

QY 1929 AACTTCAAGTGTCTCCATCAGCAAAATTTATCAGCCCGAGAGAGCAAGATTAGAGATTGA 1988
Db 61 AACTTCAAGTGTCTCCATCAGCAAAATTTATCAGCCCGAGAGAGCAAGATTAGAGATTGA 120

QY 1989 AGGAGCCCGGCTGACCTGATGAGTGTATGACATTTGAAGATTTGCTTTTAAGT 2048
Db 121 AGGAGCCCGGCTGACCTGATGAGTGTATGACATTTGAAGATTTGCTTTTGAAGT 180

QY 2049 ACAAGAGAAATGGCAAG 2108
Db 181 ACAAGAGAAATGGCAAG 240

QY 2109 TATTGACAGCAAAAG 2168
Db 241 TATTGACAGCAAAAG 300

QY 2169 TGTGCTCTCAACTCAAGAGCTTAGATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2228
Db 301 TGTGCTCTCAACTCAAGAGCTTAGATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 2229 GGTTCCTAAGGTGAGAGATAGACATGAGGTCTTATGGCTTCCTTCAAGAGAGAA 2288
Db 361 GGTTCCTAAGGTGAGAGATAGACATGAGGTCTTATGGCTTCCTTCAAGAGAGAA 420

QY 2289 GAAATATGATGAG 2348
Db 421 GAAATATGATGAG 480

QY 2349 CCCATACAGTTCGCAATGTGATGACAGAGTGGCTTCAAGAGATGATGACAGC 2408
Db 481 CCCATACAGTTCGCAATGTGATGACAGAGTGGCTTCAAGAGATGATGACAGC 540

QY 2409 TTGGCATCAAAATAGGAGCTGGCATATCTTCACCAAGAGAGAGAGAGAGAG 2468
Db 541 TTGGCATCAAAATAGGAGCTGGCATATCTTCACCAAGAGAGAGAGAGAGAG 600

QY 2469 GAAAGGCCAAG 2528
Db 601 GAAAGGCCAAG 660

QY 2529 CACAGGCTTCCTGCGCAGGAGACATCCGTTAAATATGTTCCGCCACACTGAGTCTGG 2588
Db 661 CACAGGCTTCCTGCGCAGGAGACATCCGTTAAATATGTTCCGCCACACTGAGTCTGG 720

QY 2589 AGCTATAGATGCTCATGACAGTGTGTTGACAAATGCTCCAGCCCTGAAACCTTTGTAT 2648
Db 721 AGCTATAGATGCTCATGACAGTGTGTTGACAAATGCTCCAGCCCTGAAACCTTTGTAT 780

QY 2649 TTTTAGTGGATGAGAGCTTATCTCAGTATTTGTGGACATGCCAGAGAGATTTGATCA 2708
Db 781 TTTTAGTGGATGAGAGCTTATCTCAGTATTTGTGGACATGCCAGAGAGATTTGATCA 840

QY 2709 GTCAAGAGATTTACATCATGACAGACCAATGAGACCTTTGACAGAGATCTTGGAGGAGAT 2768
Db 841 GTCAAGAGATTTACATCATGACAGACCAATGAGACCTTTGACAGAGATCTTGGAGGAGAT 900

QY 2769 CGCAGTGGGAG 2828
Db 901 CGCAGTGGGAG 960

QY 2829 CTTGGGTGAG 2888
Db 961 CTTGGGTGAG 1020

QY 2889 CATTGGTATCTGATGAG 2948
Db 1021 CATTGGTATCTGATGAG 1080

QY 2949 TAGTGCCTTATCTGCTTATGATGCTTGGGAGTGGGAGTGAATACCAATGAACAC 3008
Db 1081 TAGTGCCTTATCTGCTTATGATGCTTGGGAGTGGGAGTGAATACCAATGAACAC 1140

QY 3009 TTTTCAG 3068
Db 1141 TTTTCAG 1200

QY 3069 TATTTCATATATGAG 3128
Db 1201 TATTTCATATATGAG 1260

QY 3129 ATGATGATGATATTAAG 3188
Db 1261 ATGATGATGATATTAAG 1320

QY 3189 ATTAAGTCTTATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3243
Db 1321 ATTAAGTCTTATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1375

RESULT 7
BC003281 3012 bp mRNA linear ROD 07-AUG-2002
LOCUS BC003281
DEFINITION Mus musculus, clone MGC:7868 IMAGE:3581841, mRNA, complete cds.

D	b	1229	AACAAATACCGAGATATTGGAAGATGCAATGCAATGACGTCTGTCGTGAAATAATGGCTTTAAACA	1288
O	y	1489	AATAATACCTCCATTTTCTCTTCTGCTGCCCTTGGGACTGGAACATGGAAATAAAGAGGAA	1548
D	b	1289	GATATTAATTCATCTTCTCTTCTGCTCTCGGGACAGATGATGATTTGTAAGAGAGT	1348
O	y	1549	ACACGACAGAGATTTTGTGTGATGAAAGTTTAACTTTTGGCAAGACCATGTAAACAC	1608
D	b	1349	ACACGACGTCAGATAAATGTTTGGAGAAAGTTTGTGCAATTTGGCTAAAGACACACAGAGAAA	1408
O	y	1609	CAGTTAATCTGTAATAATTTTGGATCTTTCCACAGCAATTTGGAGATTTATATATGCTTTACGT	1668
D	b	1409	ACGCTAACTGTGAAGAATTTGTGATCTTTCCAGATGATGTGAGAACGTACAAACATTTTATAT	1468
O	y	1669	TCGTGAATGGCAAGAGAGTCCAAAGATGCTGAGTTTGGAACATTTACAGT-----GTC	1719
D	b	1469	GCTGAATGATCAAAAAAGGTGCTCAACGAGCTGGAATCTCAGGGGTAATAGTGTGCTTTAGCC	1528
O	y	1720	CCCCAGTCACACAGAGAGGAGAAAAGAAATGGGCTTGAAGCTAGATCTCTGCAATC	1779
D	b	1529	CTGAGAGTGCTCACTGGGGAGGACAAAGAAAGGCGGCTTGAAGCTTCCTCGCATC	1588
O	y	1780	AATCGATGGGATTCACAGCGGAGGAAGATGTATAGAGGCCCAAGCTGGATGCCAAAGAAATC	1839
D	b	1589	AATCTCATGGGTGTAAAAAGTGGGAGAGATGTGTAGGCCCCAGGAATGGATTTGAAGGGTTG	1648
O	y	1840	CTGAGTCTCCAGAAACCAACCATCATTTGGAATTAATCATATTTCTTACCTTGGGAGAAAG	1899
D	b	1649	CTGGCTCTCCCTGGAGCCACCAATCATTTGGAATTAATCATATTTCTTATCTTTGGGAAAAA	1708
O	y	1900	GAACTATGACATTTTGTCTGCTCAGCTTCGAAAACTTCAAGTGTCTCCATCAGCAAAATATC	1959
D	b	1709	GAGACACAGCTGTCTGTGAGCTCCAGACACACACAAAGAGTCTTCATTTTCAGAACACTGTC	1768
O	y	1960	AGCCAGGAAAGAGACAGAGTGAAGATTTGAAGAGCGCCGGGCTGACCTCATTTAGAGTGGTT	2019
D	b	1769	AGTCCAAAGAACGGCCACTTTGGAGATTTAAAGTCCCCAGGCTGACCTTATTTAGACAGATT	1828
O	y	2020	ATGAACATTTGAAGATATGCTTTGTAAAGTACAGAGAGAAATGGCAAGGAAAAAGAGCGA	2079
D	b	1829	ATGAGATTTGAATGTATGTGCTGTGACGTTTCAGGAAGAAAGTGGCAGGAAAAAGGAGAAA	1888
O	y	2080	GGCCTTTGGCGCTGCTTAGGACAGTGGACTATTACAGCACAAAAAACCCAGACGAATG	2139
D	b	1889	AATCTTTGGAGCTTGTTCAGGACAGGGGACCAACAGCAAGAAAAA---CTGATTAATAATG	1945
O	y	2140	AAACAAATAATTCATATTTTTCGAAATGTCCTGTGGCTCCACATCCAGACGACTCTGTATCA	2199
D	b	1946	GAAACAATCTGACACATTTTCACGATACCCAGCATCTTAACTTCAGGAACCTTCAGAGCCGA	2005
O	y	2200	AAGAAACAGTTTGAATAATGTGTTTGGACAGTTCCTTAAAGGTGGAGAAATGTAGCAATGAG	2259
D	b	2006	AAGAAACAGTTTGAATAATGTGTTTGGACAGTTCCTTAAAGGTGGAGAAATGTAGCAATGAG	2065
O	y	2260	GTCCTTATGGCTGCTTTCAAGAGAAAGAAATGATGGAAGAAAACTGCACAGGCCAA	2319
D	b	2066	GTCGCTGTGGCTGCTCTTCCAAAGAGAAAGAAATGATGGAAGAGAGAGGCCCAAGAGGA	2125
O	y	2320	CCGTGAGGCATATAGCGCTTTTCAGCAAGTCCCATACAGCTTCTGCAATGTGTATGACAGA	2379
D	b	2126	TCTGTGGAGCCAAAGTGTGTTTTCAGCAGGTGCCACATCATGTTCTGTCAATATACGTTGTGAGA	2185
O	y	2380	GTTGGCTTTTCAAGAAATGTACTGCAACCTTGCAGATCCAAATATACGAGAGCTGGGATTTAC	2439
D	b	2186	GTCGGCTTCCACAGAAATGTATTGTGACATCTCTTAACCCAGTTTATGAGAGCCGGCATATAT	2245
O	y	2440	TTTCAACCAAGACCTCAAAAACTGTGCAGAGAAAGGCCCAAGAAATCTCTGCTCAGATTAG	2499
D	b	2246	TTTCAACCAAGAGCCCTCAAAAAATCTAGCAGACAGCAAGGTCAAGAAAACTCTAAGCACAGACAG	2205
O	y	2500	CTGATCTATGTGTTTAGGCTGAAGTACTCACAGGCTTCTTCTGCCAGGACATCGCTTA	2559
D	b	2306	CTATCTATGTGTTTAGGCTGAAGTACTCACAGGCTTCTTCTGCCAGGAGTATTTCTCTCA	2365

TITLE
JOURNAL
COMMENT

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,R., Labouque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Lewine,J., Lieou,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McKean,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrum,J., Menes,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rotman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomas,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tittell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,W.

Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 11, 2000 this sequence version replaced g1:7283196.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L6563

Center clone name: 757-I_12

Summary Statistics

Sequencing vector: M13; M77815: 100% of reads
Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap, version 0.960731

Consensus quality: 143128 bases at least Q40

Consensus quality: 154634 bases at least Q30

Consensus quality: 159724 bases at least Q20

Insert size: 176000; agarose-fp

Quality coverage: 3.1 in Q20 bases; agarose-fp

Quality coverage: 3.3 in Q20 bases; sum-of-continfs

NOTE: This is a 'working draft' sequence. It currently
consists of 33 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

```
1 1059: contig of 1059 bp in length
* 1060 1159: gap of 100 bp
* 1160 3218: contig of 2059 bp in length
* 3219 3318: gap of 100 bp
* 3319 4939: contig of 1621 bp in length
* 4940 5039: gap of 100 bp
* 5040 5321: contig of 282 bp in length
* 5322 5421: gap of 100 bp
* 5422 6776: contig of 1355 bp in length
* 6777 6876: gap of 100 bp
* 6877 8440: contig of 1564 bp in length
* 8441 8540: gap of 100 bp
* 8541 10040: contig of 1500 bp in length
* 10041 10140: gap of 100 bp
* 10141 11902: contig of 1762 bp in length
* 11903 12002: gap of 100 bp
* 12003 15014: contig of 3012 bp in length
* 15015 15114: gap of 100 bp
* 15115 17618: contig of 2504 bp in length
* 17619 17718: gap of 100 bp
* 17719 20650: contig of 2932 bp in length
* 20651 20750: gap of 100 bp
* 20751 23090: contig of 2340 bp in length
```

```
* 23091 23190: gap of 100 bp
* 23191 26671: contig of 3481 bp in length
* 26672 26771: gap of 100 bp
* 26772 29928: contig of 3157 bp in length
* 29929 30028: gap of 100 bp
* 30029 34433: contig of 4405 bp in length
* 34434 34533: gap of 100 bp
* 34534 39381: contig of 4848 bp in length
* 39382 39481: gap of 100 bp
* 39482 44572: contig of 5091 bp in length
* 44573 44672: gap of 100 bp
* 44673 49132: contig of 4460 bp in length
* 49133 49232: gap of 100 bp
* 49233 54475: contig of 5243 bp in length
* 54476 54575: gap of 100 bp
* 54576 58555: contig of 3980 bp in length
* 58556 58655: gap of 100 bp
* 58656 64694: contig of 6039 bp in length
* 64695 64794: gap of 100 bp
* 64795 70000: contig of 5206 bp in length
* 70001 70100: gap of 100 bp
* 70101 74717: contig of 4617 bp in length
* 74718 74817: gap of 100 bp
* 74818 81551: contig of 6734 bp in length
* 81552 81651: gap of 100 bp
* 81652 87788: contig of 6137 bp in length
* 87789 87888: gap of 100 bp
* 87889 93064: contig of 5176 bp in length
* 93065 93164: gap of 100 bp
* 93165 99571: contig of 6407 bp in length
* 99572 99671: gap of 100 bp
* 99672 108010: contig of 8339 bp in length
* 108011 108110: gap of 100 bp
* 108111 116297: contig of 8187 bp in length
* 116298 116397: gap of 100 bp
* 116398 126733: contig of 10336 bp in length
* 126734 126833: gap of 100 bp
* 126834 139797: contig of 12964 bp in length
* 139798 139897: gap of 100 bp
* 139898 153345: contig of 13448 bp in length
* 153346 153445: gap of 100 bp
* 15346 167164: contig of 13719 bp in length.

FEATURES
source
1. 167164
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3"
/clone="RP11-757112"
/clone_id="RP11-757112 Human Male BAC"
1. 1059
/note="assembly_fragment"
1160. 3218
/note="assembly_fragment"
3319. 4939
/note="assembly_fragment"
5040. 5321
/note="assembly_fragment"
clone_end:77
vector_side:left"
5422. 6776
/note="assembly_fragment"
6877. 8440
/note="assembly_fragment"
8541. 10040
/note="assembly_fragment"
10141. 11902
/note="assembly_fragment"
12003. 15014
/note="assembly_fragment"
15115. 17618
/note="assembly_fragment"
17719. 20650
misc_feature
```

```
/note="assembly-fragment"
misc_feature      20751..23090
                  /note="assembly-fragment"
misc_feature      23191..26671
                  /note="assembly-fragment"
misc_feature      26772..29928
                  /note="assembly-fragment"
misc_feature      30029..34433
                  /note="assembly-fragment"
misc_feature      34534..39381
                  /note="assembly-fragment"
misc_feature      39482..44572
                  /note="assembly-fragment"
misc_feature      44673..49132
                  /note="assembly-fragment"
misc_feature      49233..54475
                  /note="assembly-fragment"
misc_feature      54576..58555
                  /note="assembly-fragment"
misc_feature      58656..64694
                  /note="assembly-fragment"
misc_feature      64795..70000
                  /note="assembly-fragment"
misc_feature      70101..74717
                  /note="assembly-fragment"
misc_feature      74818..81551
                  /note="assembly-fragment"
misc_feature      81652..87788
                  /note="assembly-fragment"
misc_feature      87889..93064
                  /note="assembly-fragment"
misc_feature      clone_end:SP6
                  vector_side:left"
```

Query Match 26.0% Score 844.4; DB 2; Length 167164;

Best Local Similarity 98.7% Pred. No.3.7e-201;

Matches 851; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
QY 380 CACAGACTGGTCTGTGGAGAAACATATGTTGGCAATTCCTCAATTAACCAATGACT 439
    |||||
DB 99684 CACAGACTGGTCTGTGGAGAAACATATGTTGGCAATTCCTCAATTAACCAATGACT 99743

QY 440 TCAAAATTTTAAAAATATAGACGCTGCTGTGTAAGTCTCCAGAAATAGTTGGCT 499
    |||||
DB 99744 TCAAAATTTTAAAAATATAGACGCTGCTGTGTAAGTCTCCAGAAATAGTTGGCT 99803

QY 500 GTATCTCTACCTGGTCTCTCCAGTTTCAGGAAGGCAACCAATCTCTGCAAGTGTCA 559
    |||||
DB 99804 GTATCTCTACCTGGTCTCTCCAGTTTCAGGAAGGCAACCAATCTCTGCAAGTGTCA 99863

QY 560 GAAAAATGCTGACTCTAGATAGAGTTATCAGTCTGGAAGATGACCTCACACACATG 619
    |||||
DB 99864 GAAAAATGCTGACTCTAGATAGAGTTATCAGTCTGGAAGATGACCTCACACACATG 99923

QY 620 CTGTGTATGCTGTGTGAATGACGCCAATGAAAGATCTTCTGCATGGGGAGGCGTGGCCC 679
    |||||
DB 99924 CTGTGTATGCTGTGTGAATGACGCCAATGAAAGATCTTCTGCATGGGGAGGCGTGGCCC 99983

QY 680 TGGCCCTGGTAAAAAGCTGGTGAATTTGAATCCAGAGAGACCAACAGTTTGGCCA 739
    |||||
DB 99984 TGGCCCTGGTAAAAAGCTGGTGAATTTGAATCCAGAGAGACCAACAGTTTGGCCA 100043

QY 740 GATATGTAAGTGTAGCTGCTGAGATAGCTTCAACGGAGAGAGGAGGCTTCCCTGCA 799
    |||||
DB 100044 GATATGTAAGTGTAGCTGCTGAGATAGCTTCAACGGAGAGAGGAGGCTTCCCTGCA 100103

QY 800 AACAGATCATCATGCTGTGGGCTCGGTGATGGAATGGATTAACAGAGATGACTG 859
    |||||
DB 100104 AACAGATCATCATGCTGTGGGCTCGGTGATGGAATGGATTAACAGAGATGACTG 100163

QY 860 GAAAGCTGACAGAGGCGCATTTGAATTTCTGAATTTATGATCATCTATAAATACTAC 919
    |||||
DB 100164 GAAAGCTGACAGAGGCGCATTTGAATTTCTGAATTTATGATCATCTATAAATACTAC 100223
```

```
QY 920 TTAAAGACAGTAGCAATTCACAGCCTTGAGCTGTGGAGATTTTTCAGTTCCCTGTAATTTCT 979
    |||||
DB 100224 TTAAAGACAGTAGCAATTCACAGCCTTGAGCTGTGGAGATTTTTCAGTTCCCTGTAATTTCT 100283

QY 980 GTACAAAGACTATTTGAGAGACTATCCGGGTGATGTTTGCAGGGAAGCCAAATGATGAGTA 1039
    |||||
DB 100284 GTACAAAGACTATTTGAGAGACTATCCGGGTGATGTTTGCAGGGAAGCCAAATGATGAGTA 100343

QY 1040 ATTTGAAAGAAATTCACCTGGTGTAGCAATGAGAGACCTCTGCTGGCTTTAAAGCTG 1099
    |||||
DB 100344 ATTTGAAAGAAATTCACCTGGTGTAGCAATGAGAGACCTCTGCTGGCTTTAAAGCTG 100403

QY 1100 CTTCAGAAATTCATCTCAGGGAAGAGTGTAGCTGTGGAGACAAGAAACACCCCTTTTCAATG 1159
    |||||
DB 100404 CTTCAGAAATTCATCTCAGGGAAGAGTGTAGCTGTGGAGACAAGAAACACCCCTTTTCAATG 100463

QY 1160 CAATGCTGTGAACAACCTGACCTCCAGATTGTCCAGGGCCACATTGAATGCGACAGCG 1219
    |||||
DB 100464 CAATGCTGTGAACAACCTGACCTCCAGATTGTCCAGGGCCACATTGAATGCGACAGCG 100523

QY 1220 CAGATGTAATTTGTAATTTCTGT 1241
    |||||
DB 100524 TAATCTTTGTTCTATTTCTCT 100545

RESULT 9
AC092908 174257 bp DNA linear PRI 16-MAY-2002
LOCUS AC092908
DEFINITION Homo sapiens 3q BAC RP11-757112 (Roswell Park Cancer Institute
ACCESSION AC092908
VERSION AC092908.9 GI:17223136
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 174257)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alstrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbraia,J.,
Benton,J., Blinag,K., Blankenburg,K., Bonini,D., Bouck,J.,
Bowle,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P., Butlay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Cartron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Doutheville,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gablis,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M.,
Holloway,C., Hollins,B., Homsli,F., Howard,S., Huber,J., Huylk,S.,
Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvath,J.,
Kovar,C., Kratovic,U., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loisaged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mahoney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,K., Miner,G., Miner,E., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenkwo,S., Ogih,M., Okuwono,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., RojudoKan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S.,
Scott,G., Shen,H., Shiohata,I., Sisson,I., Sodergren,E.,
Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
```

Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williams,A., Wlarczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and Gibbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 174257)
Worley,K.C.
Submitted (09-AUG-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 174257)
Worley,K.C.
Submitted (01-DEC-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 174257)
Worley,K.C.
Submitted (04-DEC-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 174257)
Worley,K.C.
Submitted (30-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 174257)
Worley,K.C.
Direct Submission
Submitted (16-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Dec 1, 2001 this sequence version replaced gi:17105265.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found

at URL: http://gc.bcm.tmc.edu:8088/quality_info/genbank_annotation.html.

FEATURES	QUALSTAT-REPORT.
source	Location/Qualifiers
repeat_region	1..174257 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="3" /clone="RP11-757112" complement(1446..1677) /rpt_family="MIR" complement(2119..2337) /rpt_family="MIR" 2751..2893 /rpt_family="MIR" 2963..3278 /rpt_family="AluJb" 4985..5382 /rpt_family="L2" 5397..5667 /rpt_family="AluX" complement(5672..5724) /rpt_family="MIR" complement(5678..5954) /rpt_family="L2" 6083..6228 /rpt_family="L2" complement(6330..6344) /rpt_family="MIR" complement(6345..6659) /rpt_family="AluJb" complement(6660..6848) 7058..7268 /rpt_family="MIR" 8100..8128 /rpt_family="AT-rich" 8348..8393 /rpt_family="AT-rich" 8490..8577 /rpt_family="L2" 9107..9404 /rpt_family="MTLJ2" 9405..9680 /rpt_family="AluX" 9681..9707 /rpt_family="(TAA)n" 9708..9754 /rpt_family="MTLJ2" 9895..9972 /rpt_family="MIR" 10608..10899 /rpt_family="AluJb" 13450..13752 /rpt_family="AluJc" 13762..14068 complement(14073..14217) /rpt_family="MIR" 14409..14705 /rpt_family="AluSg" 14707..14765 /rpt_family="(CCCA)n" 15196..15390 /rpt_family="MIR" 15659..15768 /rpt_family="MIR" complement(15827..15959) /rpt_family="L2" complement(17135..17446) /rpt_family="AluSc" 17725..17868

```
repeat_region      /rpt_family="MER5A"
                    18997. .19025
repeat_region      /rpt_family="(CCCCCG)n"
                    complement(20134. .20485)
repeat_region      /rpt_family="L2"
                    22275. 22626
repeat_region      /rpt_family="MER63C"
                    22627. .22650
repeat_region      /rpt_family="(TA)n"
```

Query Match	26.0%;	Score 844.4;	DB 9;	Length 174257;
Best Local Similarity	98.7%;	Pred. No. 3.7e-201;		
Matches 851;	Conservative 0;	Mismatches 11;	Indels 0;	Gaps 0;

OY	380	CAGAGACGCGGCTCTTGAGAAAACATAGTGGCAAAATTCACATTAACACAAATGACT	439
Db	144092	CAGAGACGCGGCTCTTGAGAAAACATAGTGGCAAAATTCACATTAACACAAATGACT	144151
OY	440	TCAAAATTTTAAANAATATAGCGCTGACCTGTGTGAAGTCTCCAGAAATTAAGTTGGCT	499
Db	144152	TCAAAATTTTAAANAATATAGCGCTGACCTGTGTGAAGTCTCCAGAAATTAAGTTGGCT	144211
OY	500	GTATCTTACCCGTGGTCCCTCAGCTTCAGGAAGCAACCAATCTCTCAGATGTGCA	559
Db	144212	GTATCTTACCCGTGGTCCCTCAGCTTCAGGAAGCAACCAATCTCTCAGATGTGCA	144271
OY	560	GA AAAATTCCTGACTCTTAGATATAGATTATCAGTCTGGAAGATGACCTTACCCACATG	619
Db	144272	GA AAAATTCCTGACTCTTAGATATAGATTATCAGTCTGGAAGATGACCTTACCCACATG	144331
OY	620	CTGTGTGATGCTGTGTGTGAATGACCAATGAATCTTTGTGCATGGGGAGGCTTGCC	679
Db	144332	CTGTGTGATGCTGTGTGTGAATGACCAATGAATCTTTGTGCATGGGGAGGCTTGCC	144391
-OY	680	TGGCCCTGTGTAAGCTGCTGATTTTGAATCCAAAGAAAGAGCAAAACATTTGGTGGCA	729
Db	144392	TGGCCCTGTGTAAGCTGCTGATTTTGAATCCAAAGAAAGAGCAAAACATTTGGTGGCA	144451
OY	740	GATATGTGTAAGTGTCAAGCTGTGAGATAGCTGTCAAGGGAGCAGGAGGCTTCCGTGCA	799
Db	144452	GATATGTGTAAGTGTCAAGCTGTGAGATAGCTGTCAAGGGAGCAGGAGGCTTCCGTGCA	144511
OY	800	AACAGATATCATCTGCTGTTGGGCTCGGTGTGATGGAATGGGATTAACAGGATGTACTG	859
Db	144512	AACAGATATCATCTGCTGTTGGGCTCGGTGTGATGGAATGGGATTAACAGGATGTACTG	144571
OY	860	GAAGCTGCAAGAGGCCATTGTAGTATTTGCAATTTATGTCATCTATAAANAATACACACA	919
Db	144572	GAAGCTGCAAGAGGCCATTGTAGTATTTGCAATTTATGTCATCTATAAANAATACACACA	144631
OY	920	TTTAAGACATGACATTCACGCTTGAAGCTGTGGATTTTTCAGTTCCCTGTGAATTTGT	979
Db	144632	TTTAAGACATGACATTCACGCTTGAAGCTGTGGATTTTTCAGTTCCCTGTGAATTTGT	144691
OY	980	GTACAAAGACATATTTGTAGAGACTTCCGGGTTACTTTGCAAGGGAAGCCAAATGATGATG	1039
Db	144692	GTACAAAGACATATTTGTAGAGACTTCCGGGTTACTTTGCAAGGGAAGCCAAATGATGATG	144751
OY	1040	ATTTGAAAGAAATTCACCTGTGTGACAATGAGACCCCTACTGTGTGCTGCTTTAAAGCTG	1099
Db	144752	ATTTGAAAGAAATTCACCTGTGTGACAATGAGACCCCTACTGTGTGCTGCTTTAAAGCTG	144811
OY	1100	CTTCAGAAATTCATCTCTAGGGAAGAGTGAAGCTGGGACAAGAACACCCCTTCTTTCAATG	1159
Db	144812	CTTCAGAAATTCATCTCTAGGGAAGAGTGAAGCTGGGACAAGAACACCCCTTCTTTCAATG	144871
OY	1160	CAATGTGCTGAACACACTGACCTCCAGATTTGTGCCAGGGCCCAATTTGATGTCACACGG	1219
Db	144872	CAATGTGCTGAACACACTGACCTCCAGATTTGTGCCAGGGCCCAATTTGATGTCACACGG	144931
OY	1220	CAGATGTAAATGTTTAATTTGTGT 1241	
Db	144932	TAACTCTTGTGTTTCAATTTCTCT 144953	

RESULT	10
AC026506/c	
LOCUS	167164 bp DNA linear HTG-11-APR-2000
DEFINITION	Homo sapiens chromosome 3 clone RP11-757I12 map 3, WORKING DRAFT SEQUENCE, 33 unordered pieces.
ACCESSION	AC026506.2 GI:7534005
VERSION	HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS	Homo sapiens.
SOURCE	Homo sapiens.
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 167164)
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and lander,E.
TITLE	Homo sapiens chromosome 3, clone RP11-757I12
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 167164)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Cillymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,
 Dodgson, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Glinde, S., Goyette, M., Graham, L.,
 Grand, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klewin, J., Labouque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McCurt, A., McKernan, K., McPheeters, R.,
 Melchior, J., Meneus, L., Minova, T., Miranda, C., Miengo, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Olivari, T. M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange, A., Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Stelfox, S., Theodore, J., Tirelli, A., Travers, M., Triggillo, J.,
 Vassiliev, H., Vieler, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zairoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 11, 2000 this sequence version replaced gi:1283196.

----- Genome Center
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

```
----- Project Informa
Center project name: L6563
```

```
----- Summary Statistics -----
Center clone name: 757_I_12
```

Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731
Consensus quality: 143128 bases at least Q40

Consensus quality: 154634 bases at least Q30
Consensus quality: 159724 bases at least Q20

```
Insert size: 176000; agarose-fp
Insert size: 163964; sum-of-contigs
```

quality coverage: 3.1 in Q20 bases; agarose-tf
quality coverage: 3.3 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently

consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown

his record will be updated with the finished sequence

	* as soon as it is available and the accession number will
	* be preserved.
	1 1059: contig of 1059 bp in length
*	1060 1159: gap of 100 bp
*	1160 3218: contig of 2059 bp in length
*	3219 3318: gap of 100 bp
*	3319 4939: contig of 1621 bp in length
*	4940 5039: gap of 100 bp
*	5040 5321: contig of 282 bp in length
*	5322 5421: gap of 100 bp
*	5422 6776: contig of 1355 bp in length
*	6777 6876: gap of 100 bp
*	6877 8440: contig of 1564 bp in length
*	8441 8540: gap of 100 bp
*	8541 10040: contig of 1500 bp in length
*	10041 10140: gap of 100 bp
*	10141 11902: contig of 1762 bp in length
*	11903 12002: gap of 100 bp
*	12003 15014: contig of 3012 bp in length
*	15015 15114: gap of 100 bp
*	15115 17618: contig of 2504 bp in length
*	17619 17718: gap of 100 bp
*	17719 20650: contig of 2932 bp in length
*	20651 20750: gap of 100 bp
*	20751 23090: contig of 2340 bp in length
*	23091 23190: gap of 100 bp
*	23191 26671: contig of 3481 bp in length
*	26672 26771: gap of 100 bp
*	26772 29928: contig of 3157 bp in length
*	29929 30028: gap of 100 bp
*	30029 34433: contig of 4405 bp in length
*	34434 34533: gap of 100 bp
*	34534 39381: contig of 4848 bp in length
*	39382 39481: gap of 100 bp
*	39482 44572: contig of 5091 bp in length
*	44573 44672: gap of 100 bp
*	44673 49132: contig of 4460 bp in length
*	49133 49232: gap of 100 bp
*	49233 54475: contig of 5243 bp in length
*	54476 54575: gap of 100 bp
*	54576 58555: contig of 3980 bp in length
*	58556 58655: gap of 100 bp
*	58656 64694: contig of 6039 bp in length
*	64695 64794: gap of 100 bp
*	64795 70000: contig of 5206 bp in length
*	70001 70100: gap of 100 bp
*	70101 74717: contig of 4617 bp in length
*	74718 74817: gap of 100 bp
*	74818 81551: contig of 6734 bp in length
*	81552 81651: gap of 100 bp
*	81652 87788: contig of 6137 bp in length
*	87789 87888: gap of 100 bp
*	87889 93064: contig of 5176 bp in length
*	93065 93164: gap of 100 bp
*	93165 99571: contig of 6407 bp in length
*	99572 99671: gap of 100 bp
*	99672 108010: contig of 8339 bp in length
*	108011 108110: gap of 100 bp
*	108111 116297: contig of 8187 bp in length
*	116298 116397: gap of 100 bp
*	116398 126733: contig of 10336 bp in length
*	126734 126833: gap of 100 bp
*	126834 139797: contig of 12964 bp in length
*	139798 139897: gap of 100 bp
*	139898 153345: contig of 13448 bp in length
*	153346 153445: gap of 100 bp
*	153446 167164: contig of 13719 bp in length.
	Location/Qualifiers
	1..167164
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="3"
	/map="3"

	/clone="RP11-757112"
	/clone_lib="RPC1-11 Human Male BAC"
	1..1059
	/note="assembly-fragment"
	1160..3218
	/note="assembly-fragment"
	3319..4939
	/note="assembly-fragment"
	5040..5321
	/note="assembly-fragment"
	clone_end:77
	vector_side:left
	5422..6776
	/note="assembly-fragment"
	6877..8440
	/note="assembly-fragment"
	8541..10040
	/note="assembly-fragment"
	10141..11902
	/note="assembly-fragment"
	12003..15014
	/note="assembly-fragment"
	15115..17618
	/note="assembly-fragment"
	17719..20650
	/note="assembly-fragment"
	20751..23090
	/note="assembly-fragment"
	23191..26671
	/note="assembly-fragment"
	26772..29928
	/note="assembly-fragment"
	30029..34433
	/note="assembly-fragment"
	34534..39381
	/note="assembly-fragment"
	39482..44572
	/note="assembly-fragment"
	44673..49132
	/note="assembly-fragment"
	49233..54475
	/note="assembly-fragment"
	54576..58555
	/note="assembly-fragment"
	58656..64694
	/note="assembly-fragment"
	64795..70000
	/note="assembly-fragment"
	70101..74717
	/note="assembly-fragment"
	74818..81551
	/note="assembly-fragment"
	81652..87788
	/note="assembly-fragment"
	87889..93064
	/note="assembly-fragment"
	clone_end:SP6
	vector_side:left

	Query Match
	Best Local Similarity 100.0%; Pred. No. 2.9e+19s;
	Matches 821; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	2413 GATCAAAATCGGAGCTGCATATTACTTCCACAAGACCTCAAAAACCTGGCAGAGAAG 2472
Dd	90675 GATCAAAATTCGAGCTGCATATTACTTCCACAAGACCTCAAAAACCTGGCAGAGAAG 90616
Oy	2473 GCCAAGAAAAATCTGCTGCAGATAAGCTGATGTGTTGAGGCTGAAGTAAGTACACA 2532
Dd	90615 GCCAAGAAAAATCTGCTGCAGATAAGCTGATGTGTTGAGGCTGAAGTAAGTACACA 90556
Oy	2533 GCCTTTCTTGCCAGGAGCATCCGTTTAATATTGTTCCGCCACACTGAGTCCGTGAGCT 2592

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoideae; Homo.									
AUTHORS	1 (bases 1 to 666)									
TITLE	Schlegel, R., Deeds, J., Berger, A. and Zhao, X. Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer									
JOURNAL	Patent: WO 0142467-A 3936 14-JUN-2001; Millennium Predictive Medicine, Inc. (US)									
FEATURES	Location/Qualifiers									
SOURCE	1..666									
BASE COUNT	/organism="Homo sapiens"									
ORIGIN	/db_xref="taxon:9606"									
Query Match	15.1%; Score 489.4; DB 6; Length 666;									
Best Local Similarity	84.6%; Pred. No. 5.4e-112;									
Matches 608; Conservative	0; Indels 105; Gaps 1;									
QY	136	GCAGAAAGTGGGCGACCATATCTGGAAGACTATACAGTATGCTTTGAAGGCGCAAAAGGAAAT	195							
Db	52	GAAAAAGTGGGCGCCACCATATCTGGAAGACTATACGTTATGCTTTGAAGGCGCAAAAGGAAAT	111							
QY	196	AAACATTTAAAGACTCCCGCGGAGCACTGGAAGATGACCTTTCCATGGTGGCGCGAGACA	255							
Db	112	AAACATTTAAAGACTCCCGCGGAGCACTGGAAGATGACCTTTCCATGGTGGCGCGAGACA	171							
QY	256	GCAGCTTCATGAAAATATCAGTAGATTTACCTCGCTCTCACTCTGTTTCAGAAAGTC	315							
Db	172	GCAGCTTCATGAAAATATCAGTAGATTTACCTCGCTCTCACTCTGTTTCAGAAAGTC	189							
QY	316	TTTGCTCAGATCTTCTCAGTAGGAGAAAGGGAATACAGAAAGATGTCTCCCTACAG	375							
Db	190	-----	189							
QY	376	TGCTCAGAGACTGCTGCTCTTGAGAAAACATATAGTTGGCAATTCCTTAACCAAT	435							
Db	190	---TCAGAGACTGCTGCTCTTGAGAAAACATATAGTTGGCAATTCCTTAACCAAT	246							
QY	436	GACTTCAAAATTTTAAAAATATAGCGCTCAGCTGTGTGAAGTCTCCAGAAATTAATT	495							
Db	247	GACTTCAAAATTTTAAAAATATAGCGCTCAGCTGTGTGAAGTCTCCAGAAATTAATT	306							
QY	496	GCGCTATCTCACCCTGGCTCTCCACATTCGAGAAAGCAACAGAAATCTCTGCAAGT	555							
Db	307	GCGCTATCTCACCCTGGCTCTCTCCAGTTCGAGAAAGCAACAGAAATCTCTGCAAGT	366							
QY	556	TTCAAAAAAATCTGACTCTCTAGATAGATATAGTCTGGAAGATGACTCACCA	615							
Db	367	TTAACAAAAATCTGACTCTCTAGATAGATATAGTCTGGAAGATGACTCACCA	426							
QY	616	CATGCTGTTGATGCTGTGGTGAATGCAGCCCAATGAAGATCTTCTCATGGGGAGGCGCTG	675							
Db	427	CATGCTGTTGATGCTGTGGTGAATGCAGCCCAATGAAGATCTTCTCATGGGGAGGCGCTG	486							
QY	676	GCCCTGGCCCTGGTAAAAAGCTGTGATTGAAATCCAAAGAGACGAAACAGTTGTT	735							
Db	487	GCCCTGGCCCTGGTAAAAAGCTGTGATTGAAATCCAAAGAGACGAAACAGTTGTT	546							
QY	736	GCCAAATATGGAAGTGTGACGTGTGATGAGCTGTACCGGAGACAGGAGGCTTCCC	795							
Db	547	GCCAAATATGGAAGTGTGACGTGTGATGAGCTGTACCGGAGACAGGAGGCTTCCC	606							
QY	796	TGCAACAAGATATCCATGCTTTGGGCTCGGTGATGGAATGGAATTAACAGGATG	854							
Db	607	TGCAACAAGATATCCATGCTTTGGGCTCGGTGATGGAATGGAATTAACAGGATG	665							
RESULT 13										
LOCUS	AX185819/c 516 bp DNA Linear PAT 06-AUG-2001									
DEFINITION	Sequence 1514 from Patent WO0142467.									
ACCESSION	AX185819									
VERSION	AX185819.1 GI:15137230									

KEYWORDS	human.									
SOURCE	Homo sapiens									
ORGANISM	Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumetazoa; Euteleostomi; Primates; Catarrhini; Hominiidae; Homo.									
REFERENCE	1 (bases 1 to 516)									
AUTHORS	Schlegel, R., Deeds, J., Berger, A. and Zhao, X.									
TITLE	Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer									
JOURNAL	Patent: WO 0142467-A 1514 14-JUN-2001; Millennium Predictive Medicine, Inc. (US)									
FEATURES	Location/Qualifiers									
source	1..516									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
BASE COUNT	124 a 138 c 108 g 146 t									
ORIGIN										
Query Match	14.3%: Score 463.2; DB 6; Length 516;									
Best Local Similarity	99.2%: Pred. No. 2.1e-105;									
Matches	476; Conservative 0; Mismatches 3; Indels 1; Gaps 1;									
QY	379	TCAGAGACTGTCGTCCTTGAGAAACATATGTTGGCAA--TTCCCATTTAAACAAATGA	437							
DB	507	TCAGAGATTGGTGCTTGAGAAACATATGTTGGCAAATTTCCCATTTAAACAAATGA	448							
QY	438	CTTCAAAATTTTAAAAATATGACGCTGCTGTGAAGTCTCTCGAATAGTTGG	497							
DB	447	CTTCAAAATTTTAAAAATATGACGCTGCTGTGAAGTCTCTCGAATAGTTGG	388							
QY	498	CTGTATCTCTACCCGTCCTCCAGTTGAGAAAGCAACGCAATCTGCAAGTGT	557							
DB	387	CTGTATCTCTACCCGTCCTCCAGTTGAGAAAGCAACGCAATCTGCAAGTGT	328							
QY	558	CAGAAAAATGCTGACCTCTAGAGATAGAGTTATCACTGTGAAAGATGACCTCACACACA	617							
DB	327	CAGAAAAATGCTGACCTCTAGAGATAGAGTTATCACTGTGAAAGATGACCTCACACACA	268							
QY	618	TGCTGTGATGCTGTGTGAATGACGCCAATGAGATCTTCTGCATGGGGAGGCTTGGC	677							
DB	267	TGCTGTGATGCTGTGTGAATGACGCCAATGAGATCTTCTGCATGGGGAGGCTTGGC	208							
QY	678	CCTGGCCCTGTTAAAGCTGTGTGAATTTGAAATCCAAAGAGAGCAAACTTTGTTGC	737							
DB	207	CCTGGCCCTGTTAAAGCTGTGTGAATTTGAAATCCAAAGAGAGCAAACTTTGTTGC	148							
QY	738	CAGATATGTTAAAGTGTCAGCTGGAGATAGCTCTACGGGAGGAGGAGGCTTCCCTG	797							
DB	147	CAGATATGTTAAAGTGTCAGCTGGAGATAGCTCTACGGGAGGAGGAGGCTTCCCTG	88							
QY	798	CAAACAGATCATCATGCTGTTGGGCTCGGTGATGAGATGGGATGAAACAGGATGTAC	857							
DB	87	CAAACAGATCATCATGCTGTTGGGCTCGGTGATGAGATGGGATGAAACAGGATGTAC	28							
RESULT 14										
LOCUS	AX187250/c	481 bp	DNA	linear	PAT 06-AUG-2001					
DEFINITION	Sequence 2945 from Patent W00142467.									
ACCESSION	AX187250									
VERSION	AX187250.1 GI:15138702									
KEYWORDS	human.									
SOURCE	Homo sapiens									
ORGANISM	Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumetazoa; Euteleostomi; Primates; Catarrhini; Hominiidae; Homo.									
REFERENCE	1 (bases 1 to 481)									
AUTHORS	Schlegel, R., Deeds, J., Berger, A. and Zhao, X.									
TITLE	Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer									
JOURNAL	Patent: WO 0142467-A 2945 14-JUN-2001; Millennium Predictive Medicine, Inc. (US)									
FEATURES	Location/Qualifiers									

